Spatial omics in neuronal cells:

What goes where and why?

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

von David van den Bruck

Eingereicht an der Lebenswissenschaftlichen Fakultät der Humboldt Universität zu Berlin

Präsidentin der Humboldt Universität zu Berlin: Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät: Prof. Dr. Bernhard Grimm

Gutachter: 1. Prof. Dr. Christian Schmitz-Linneweber

2. Dr. Marina Chekulaeva

3. Prof. Dr. Matthias Selbach

Tag der mündlichen Prüfung: 17.06.2019

Erklärung über die selbstständige Abfassung meiner Dissertation

Hiermit erkläre ich, David van den Bruck, Einschreibnummer: 565072, dass ich die vorliegende Dissertation selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.
Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Während der Entstehung dieser Dissertation wurde mit keinem gewerblichen Promotionsberater zusammengearbeitet.
Die dieser Dissertation zugrunde liegende Arbeit wurde am Max-Delbrück-Centrum für Molekulare Medizin in der Arbeitsgruppe Chekulaeva in den Jahren 2014 bis 2017 getätigt. Die Dissertation selbst wurde in den anschließenden sechs Monaten bis einschließlich Juni 2018 verfasst. Später erschienene Publikationen zu diesem Thema konnten leider nicht mehr berücksichtigt werden.
Die Dissertation wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt oder veröffentlicht. Ebenso versichere ich, dass ich mich bisher nicht anderweitig um einen Doktorgrad beworben habe oder einen besitze. Ich habe die dem Promotionsverfahren zugrundeliegende Promotionsordnung zur Kenntnis genommen und versichere, dass ich die Grundsätze der Humboldt Universität zur Sicherung guter wissenschaftlicher Praxis eingehalten habe.

Kleve, den 01.01.2019

Contents

I. ABSTRACT 7

II. LIST OF FIGURES 9

III. ABBREVIATIONS 10

IV. ACKNOWLEDGEMENT 13

1. INTRODUCTION 14

1.1. Development of multicellular organisms and the importance of molecular trafficking 14

1.1.1. Cell proliferation: The cell cycle 14

1.1.2. Cell-cell interactions: Cell junctions and the extracellular matrix (ECM) 17

1.1.3. Cell specialization: Stem cells, differentiation and tissue renewal 18

1.1.4. Cell movement: Internal structure and organization of the mammalian cell 20

1.2. Mechanisms and importance of cellular polarity 23

1.2.1. Cell polarity in neuronal plasticity 25

1.2.2. In the wrong place at the wrong time: Cellular polarity in development, disease and cancer 29

1.3. The “triple T” destiny of an mRNA: Transcription, Transport, Translation 31

1.3.1. Transcription: The birth of a new RNA 31

1.3.1.1. Splicing and alternative splicing 32

1.3.1.2. Capping and poly-adenylation 34

1.3.2. Importance and mechanisms of RNA localization 35

1.3.2.1. Proteins that bind and manipulate RNAs in versatile ways (RBPs) 36

1.3.2.2. Zip codes: RNA sequences that regulate mRNA localization 37

1.3.3. Spatial translational control: Mechanisms of posttranscriptional gene regulation 41

1.3.3.1. RNA editing 41

1.3.3.2. Nonsense-mediated mRNA decay (NMD) 42

1.3.3.3. RNA regulation by noncoding RNAs: miRNAs, siRNAs, piRNAs, snoRNAs and lncRNAs 43

1.3.3.4. mRNA m6A methylation: Writers, readers and erasers 45

1.3.3.5. Translational initiation and cap-binding proteins 47

1.3.3.6. mRNA structure and stability 48

1.3.3.7. uORFs, alternative ORFs and overlapping ORFs 49

1.4. Methodical evolution in studying RNA localization in neurons: From the FISH to the mouse TRAP and beyond 50

1.4.1. In vivo systems to study intracellular localization in neurons 52
1.4.1.1. Microdissection followed by sucrose, Ficoll or Percoll density gradient centrifugation 52
1.4.1.2. Microdissection coupled to NGS 53
1.4.1.3. Microdissection followed by affinity purification techniques coupled to NGS 54
1.4.2. Cell culture-based systems and their usage in investigating RNA localization 56
1.4.2.1. Microfluidic chambers 59
1.4.2.2. Campenot chamber 59
1.4.2.3. Modified Boyden chamber 59
2. AUTHOR CONTRIBUTION 61
3. AIM OF THE THESIS 62
4. MATERIAL AND METHODS 63
4.1. Material 63
4.1.1. Equipment and consumables 63
4.1.1.1. Consumables for cell culture 63
4.1.1.2. Consumables for biochemistry assays 63
4.1.1.3. Equipment 65
4.1.1.4. Analysis – software 65
4.1.2. Chemicals and enzymes 66
4.1.3. Kits 68
4.1.4. Buffers, solutions and media 70
4.1.5. Oligos 72
4.1.6. Plasmids 90
4.2. Methods 90
4.2.1. Cell culture 90
4.2.1.1. Culture and differentiation of N1E-115 90
4.2.1.2. Culture and differentiation of Ascl1-mESC 91
4.2.1.3. RNA transfection 92
4.2.1.4. DNA transfection 92
4.2.1.5. Generation of stable cell lines using the piggyback transposon system 92
4.2.1.6. Separation of subcellular compartments of N1E-115 92
4.2.1.7. Separation of subcellular compartments of Ascl1-iNs 93
4.2.2. Biochemical assays 94
4.2.2.1. Cloning 94

4.2.2.2. Preparation of samples for liquid chromatography tandem mass spectrometry (LS MS/MS) 95
4.2.2.3. Stable isotope labelling by amino acids (SILAC) 95
4.2.2.4. Pulsed stable isotope labelling by amino acids (pSILAC) 96
4.2.2.5. Quantitative bio-orthogonal noncanonical amino acid tagging with stable-isotope labelling by amino acids (QuaNCAT) 97
4.2.2.6. SDS-PAGE 98
4.2.2.7. Western blotting 101
4.2.2.8. RNA extraction 101
4.2.2.9. cDNA synthesis and quantitative real time PCR (qRT-PCR) 102
4.2.2.10. NanoString 102
4.2.2.11. Total RiboZero RNA sequencing 102
4.2.2.12. Small RNA sequencing 103
4.2.2.13. In vitro transcription and A-tailing 103
4.2.2.14. Modified RNA element selection assay (RESA) 103
4.2.3. Microscopy 106
4.2.3.1. Immunohistochemistry 106
5. RESULTS 109
5.1. Characterization and differentiation of the cell line N1E-115 109
5.2. Characterization and differentiation of Ascl-mESC into Ascl1-iNs 111
5.3. The local transcriptome and proteome of neuronal cells revealed by spatial omics 113
5.3.1. Local -omics analysis of the neuroblastoma cell line N1E-115 114
5.3.2. Local –omics analysis of Ascl1 induced neurons derived from mouse embryonic stem cells 116
5.3.3. The local transcriptome and the local proteome is cell line specific 120
5.3.4. Identification of differential isoform localization, lncRNAs and circRNA localization 121
5.3.5. RNA localization correlates with protein localization in Ascl1 derived neurons 124
5.3.6. Overlay with spatial ribosome profiling data reveals the impact of local translation on the spatial proteome 126
5.3.7. Pools of localized RBPs and miRs as potential players of RNA localization and regulators of local translation 130
5.4. Going postal: Towards a zip-code directory for neuronal cells 134
5.4.1. Motive enrichment analysis of neurite enriched RNAs 134
5.4.2. Using an RNA reporter library to identify zip-codes in Ascl1 iNs 135
5.4.3. Using a transgenic reporter cell line to identify zip-codes in Ascl1-Ins 139

5.4.4. The RNA element selection assay (RESA) provides the tools for precise mapping of cis-regulatory elements of RNA localization 141
5.4.5. Generation of a plasmid-based RESA library 144
6. DISCUSSION 146
6.1. The neuronal model systems N1E-115 and Ascl1-iNs 146
6.2. The role of RNA localization and local translation in establishing local protein pools in neurons 147
6.3. Spatial translational regulation in outgrowing neurites 152
6.4. Spatial circular RNAs bear the potential for unknown functions in subcellular compartments of neurons 156
6.5. Identification of cis-regulatory elements and trans acting factors of RNA localization in Ascl1 iNs 157
6.5.1. Problems and solutions of the RESA-approach 157
7. REFERENCES 158

I. Abstract

Intracellular protein and RNA localization is one of the mayor players in the formation of cell shape, enabling cell agility, cellular differentiation and cell signaling. Various diseases are associated with malfunctions of intracellular molecule transport. There are many known pathways of how and why proteins and RNAs are transported within the cell and where they are located, though there is not much known about the global distribution of proteins and RNAs within the cell.

In this study I apply a subcellular fractionation method coupled to multiple omics approaches to investigate the global distribution of mRNAs, noncoding RNAs and proteins in neuronal cells. Neurites and soma from mouse neuroblastoma cells (N1E-115) as well as from Ascl1 induced neurons (Ascl1-iNs) were isolated and the composition of the spatial proteome and transcriptome was examined.

The localization of mRNAs correlates significantly with the localization of their corresponding protein products in Ascl1-iNs whereas it does not in the mouse neuroblastoma cell line N1E-115. Comparing these datasets with recently published data of other cell lines and methods it is clear, that the local proteome, transcriptome and translatome of neuronal cells is highly cell type specific.

To investigate how spatial protein pools are established I analyzed local pools of newly synthesized proteins revealing that many proteins are synthesized on the spot. RNA localization therefore plays a crucial role in generating local protein pools in these highly polarized cell systems.

Additionally, I propose a method to identify on a global scale *de novo* "zip codes" in these cell systems which would be a great step towards understanding malfunctions in molecule transport.

I. Zusammenfassung

Intrazelluläre Protein- und RNA-Lokalisation ist ein lebenswichtiger molekularer Mechanismus. Ihm unterliegen sowohl die äußere Gestaltung der Zellform, Zellagilität, zelluläre Differenzierung sowie die intra- sowie interzelluläre Kommunikation. Diverse Krankheiten werden mit Fehlfunktionen des intrazellulären Molekültransportes assoziiert und es existieren unzählige Beispiele für bekannte Wege des intrazellulären Protein- und RNA-Transportes. Allerdings ist die globale Komposition lokaler Protein- und RNA-Reservoirs bisher kaum wissenschaftlich erforscht worden.

In dieser Studie beschreibe ich die Protein- sowie RNA-Kompositionen subzellulärer Fraktionen zweier neuronaler Zelltypen. Die Neuriten und Somata von Neuroblastoma-Zellen (N1E-115) und Ascl1 induzierten Neuronen (beides Mauszellen) wurden mechanisch voneinander separiert und mittels RNA-Sequenzierung und Massenspektrometrie auf ihre Bestandteile untersucht.

Die Verteilung von mRNAs korreliert signifikant mit der Verteilung der entsprechenden Proteine in Ascl1-iNs während in der Neuroblastoma Zelllinie N1E-115 kein solcher Trend nachgewiesen werden konnte. Der Vergleich zu Datensätzen von anderen Zellsystemen und Methoden zeigt, dass das lokale Proteom sowie das lokale Transkriptome und Translatome stark Zelltyp spezifisch ist.

Um den Einfluss lokaler Proteinbiosynthese auf die Komposition subzellulärer Proteinpools zu erheben, habe ich die Lokalisation neu synthetisierter Proteine untersucht. Es scheint, als sei die RNA-Lokalisation und lokale Translation von hoher Relevanz für die Protein-Lokalisation in diesen stark polarisierten Zellsystemen.

Des Weiteren stelle ich eine Methode vor, um *de novo* „zip codes“ in diesen neuronalen Zellsystemen zu identifizieren. Diese könnte ein elementar wichtiger Schritt sein, um Fehlfunktionen im interzellulären Molekültransport zu verstehen.

II. List of Figures

- Fig. 1 Cell junctions
- Fig. 2 The cytoskeleton
- Fig. 3 Mechanisms of mRNA localization
- Fig. 4 Pulsed Stable Isotope Labelling by Amino Acids (pSILAC)
- Fig. 5 Quantitative bio-orthogonal noncanonical amino acid tagging with stable isotope labelling of amino acids (QuaNCAT)
- Fig. 6 Fragmentation and gel purification of RESA fragments
- Fig. 7 Cloning of RESA fragments
- Fig. 8 Characterization of the neuroblastoma cell line N1E-115
- Fig. 9 Characterization of Ascl1-iNs
- Fig. 10 Schematic presentation of the separation of subcellular compartments
- Fig. 11 Subcellular compartment separation of the neuroblastoma cell line N1E-115
- Fig. 12 Subcellular compartment separation followed by proteome analysis of Ascl-iNs
- Fig. 13 Transcriptome analysis of subcellular compartments of Ascl1-iNs
- Fig. 14 Comparison of the neuronal model systems N1E-115 and Ascl1-iNs
- Fig. 15 Differential Exon localization in Ascl1-iNs
- Fig. 16 Comparison of transcript localization and corresponding protein products in neuronal model systems
- Fig. 17 Local translation in neurites of Ascl1-iNs
- Fig. 18 Validation of local translation
- Fig. 19 Spatial miRs in the neuronal model systems N1E-115 and Ascl1-iNs
- Fig. 20 RBP localization in the used model systems
- Fig. 21 Targetome of either neuritic or somatic enriched miRs in Ascl1-iNs
- Fig. 22 Motive enrichment analysis of localized transcripts
- Fig. 23 Schematic presentation of high throughput identification of zip codes in Ascl1-iNs
- Fig. 24 Employing a total and transfected RNA reporter library in Ascl1-iNs
- Fig. 25 Employing a polyclonal reporter cell line to identify zip codes in Ascl1-iNs
- Fig. 26 Motive enrichment analysis of localized sequences
- Fig. 27 Customizing the RESA protocol to identify zip-codes in neurons
- Fig. 28 Diagnostics of the plasmid-based RESA experiment
- Fig. 29 Our data in the context of current literature
- Fig. 30 Spatial translational regulation is the main regulator of the local proteome

III. Abbreviations

ABCP	apico-basal cell polarity
ACD	asymetric cell division
AHA	azidohomoalanine
AIS	axon initial segment
ALS	amyotrophic lateral sclerosis
AMIS	apical membrane initiation site
APC	anaphase promoting complex
APS	ammonium persulfate
BRE	tfiib binding element
circRNA	circular rna
CKI	cyclin-dependent inhibitory protein
CLIP	cross linking coupled to immunoprecipitation
CNS	central nerve system
CTD	c-terminal domain
DAPI	4′,6-diamidin-2-phenylindol
DEPC	diethyl pyruvate
DMEM	dulbeccos modified eagle medium
DMSO	dimethyl sulfoxide
DPE	downstream promotor element
ECM	extracellular membrane
EJC	exon junction complex
EMT	epithelial-mesenchymal transition
ER	endoplasmic reticulum
ERM protein	ezrin, radixin and moesin proteins
ESC	embryonic stem cells
ESE	exonic splice enhancer
es-FBS	fetal bovine serum – suitable for embryonic stem cells
ESS	exonic splice silencer
FBS	fetal bovine serum
FISH	fluorescent *in situ* hybridization
FragSeq	fragmentation sequencing
FRAP	fluorescent recovery after photobleaching
FRCP	front-rear cell polarity
GEF	guanine exchange factor
GFP	green fluorescent protein
GO	gene ontology
IHC	immunohistochemistry
iNs	induced neurons
IP	immunoprecipitation
iPSC	induced pluripotent stem cell
IR	insulin receptor
IRES	internal ribosomal entry site
ISH	*in situ* hybridization
ISS	intronic spice silencer
KH domain	k-homology domain
LB	liquid broth
lncRNA	long noncoding RNA
LPA	lysophosphatidic acid

LS	liquid chromatography
LTP	long term potentiation
LUHMES	Lund human mesencephalic cells
mESC	mouse embryonic stem cells
miRNA	micro RNA
MISO	Mixture of isoforms (probabilistic model)
MS	mass spectrometry
MTE	motif ten element
NGS	next generation sequencing
NMD	nonsense mediated decay
ORC	origin recognition complex
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PAR proteins	partitioning defective proteins
PARS	parallel analysis of RNA structure
P-body	processing body
PBS	phosphate buffered saline
PCP	planar cell polarity
PCR	polymerase chain reaction
PIP	phosphatidylinositol-bisphosphate
piRNA	piwi-interacting RNA
PRC	polycomb repressive complex
preRC	pre-replicatory complex
PRP	plasticity related products
PSI	percentage sliced in
pSILAC	pulsed stable isotope labelling by amino acids
PVDF	polyvenyl difluoride
qRT-PCR	quantitative real-time polymerase chain reaction
QuanCAT	quantitative bio-orthogonal noncanonical amino acid tagging with stable isotope labelling by amino acids
RBP	RNA-binding protein
RESA	RNA element selection assay
RGC	retinal ganglion cells
RIP	RNA immunoprecipitation
RISC	RNA induced silencing complex
RNA-seq	RNA sequencing
RNP	ribonucleoprotein
RRM	RNA recognition motif
SDS	sodium dodecyl sulfate
SHAPE	selective 2-hydroxyl acylation analyzed by primer extension
SILAC	stable isotope labelling by amino acids
siRNA	small interfering RNA
SMA	spinal muscular atrophy
SMN	survival of motor neuron protein
TBE	tris borate EDTA
TBS	tris buffered saline solution
TC	ternary complex
TEMED	tetramethylethylenediamine
TRAP	translating ribosome affinity purification

uORF	upstream open reading frame
UTR	untranslated region

IV. Acknowledgement

Firstly, I want to thank my fellow labmates Alessandra Zappulo, Camilla Cioli Mattioli, Larissa Ruhe, Dr. Esteban Peguero Sanchez and Dr. Marta Mauri for the fruitful discussions, moral support and pleasant atmosphere during these past long years of research and the writing of this thesis.

I would also like to thank our neighbors from the Zinzen group Dr. Robert Zinzen, Philipp Wahle and Alexander Glahs. Cristal Peck needs to be mentioned here too for providing additional support wherever it was needed.

Special thanks also go to my collaborators, friends and advisors, Dr. Vedran Franke, Dr. Koshi Imami, Dr. Lorenzo Calviello, Dr. Erik McShane, Dr. Andrei Filipchyk and Dr. Guido Mastrobuoni, and to their superiors Dr. Altuna Akalin, Prof. Uwe Ohler, Prof. Mathias Selbach, Prof. Nikolaus Rajewsky and Dr. Stefan Kempa for providing this awesome scientific community.

I would like to express gratitude to my PhD committee members Dr. Stefan Kempa, Prof. Wei Chen, Dr. Jan Philipp Junker and of course Dr. Marina Chekulaeva for their continuous advice and support during the course of my work.

For continuous moral support and distraction when it was needed, I thank my wife, Wiebke and my son, Leonard.

Last but not least, I want to express sincere thanks to my doctor father Prof. Christian Schmitz-Linneweber for reviewing and supporting my thesis and of course my supervisor Dr. Marina Chekulaeva for providing the space, funding and opportunity for this work and for the long and, at times, tedious hours of supervision over the past few years!

1. Introduction

1.1. Development of multicellular organisms and the importance of molecular trafficking

Most animals and plants start as single cells, but they do have very different and distinct ways to form multicellular organisms. This process is depicted broadly as a self–assembly process with four fundamental forces: cell proliferation, cell-cell interactions, cellular differentiation and cell movement. All these conserved mechanisms shape the animal body plan starting from a single fertilized zygote. This unique cell carries a lot of maternal information and resources and cleaves rapidly without growing. It is by definition omnipotent, since it carries the potential to form each cell of the organism. During maternal to zygotic transition the blastula is formed and embryonic development can be observed as we know it from textbooks [1]. External cells form the ectoderm, invaginated cells form the endoderm and cells moving in between these two layers form the mesoderm [2]. At this point most cells are pluripotent. During further development the cells become increasingly restricted in their potential for differentiation. This process reaches its limits when cells undergo terminal differentiation. The fate of a cell relies not only on its genetic material but also on a combination of endogenous genetic information, external factors and historical burdens – the cell memory [3], [4]. Cell-cell contacts, even if they are transient and weak, greatly impact the future of both cells [5] and genes involved in cell-cell communication, transcription, while chromatin structure and regulatory DNA are of crucial importance during all steps of embryonic development [6]. Through combinatorial control and cell memory even simple plans can form complex patterns [7]. The first steps of this developmental patterning can be observed within the first seconds of embryonic life: polarization of the embryo before gastrulation and asymmetric cell divisions generate cell diversity within the very first moments of all of our lives [8], [9]. Throughout this process three major axes are generally formed: the animal-vegetal axis, which defines whether parts of the body are ex- or internal; the anterior-posterior axis, which defines the future head-tail orientation of the organism; and the dorso-ventral axis, which specifies the future back and belly orientation. At one extreme – the insect egg – these axes are already defined by the outer appearance of the zygote. At another extreme – the mammalian development – the egg is spherical, and the axis must be formed from inside. A combination of signaling proteins patterns the vertebrate embryo and forms the organism as we know it [10].

1.1.1. Cell proliferation: The cell cycle

The only way to from a new cell is to duplicate an existing cell. In unicellular species the division results in two new organisms, whereas multicellular species are a result of a cascade

of cell divisions. Continuous cell divisions are crucial for the sustainability of higher organisms. The two main tasks are (1) dividing the cell body and meanwhile (2) passing on the genetic code to the two new daughter cells. This process is called the cell-cycle and can be divided into four phases: G1, S, G2 and M-phase. The event is controlled by a number of feedback loops and irreversible ON/OFF switches and results in a duplication of the genetic material (mitosis), cell organelles and cytoplasmic division (cytokinesis) [11]. The control system of the cell cycle operates like a timer and is remarkably robust and highly adjustable. It consists of a cascade of cyclically expressed and activated cyclin-dependent protein kinases (Cdks) and blocks progression if, at any step, a mistake is detected. The activities of these kinases rise and fall as the cell progresses through the cycle and with them, the phosphorylation of intracellular protein groups that initiate or regulate the process changes. Cyclins, which are proteins that are expressed and degraded in a cyclic dependent manner, bind these kinases and inhibit their activity [12]. There are four classes of cyclins, three of which are required for all eukaryotic cells: G1/S-cyclin, which activate Cdks in late G1 phase an initiate the cycle; S-cyclins, which bind Cdks soon after progression and help to stimulate chromosome duplication; and M-cyclins, which activate Cdks that stimulate mitosis and G1-cyclins that help in most cells to govern the G1/S cyclin work [13]. In vertebrate cells there are additionally four classes of Cdks: two of these classes interact with G1-cyclins, one with G1/S-cyclins and one with S- and M-cyclins. Cyclins do not only activate the kinase activity of Cdks but also help to recognize their targets. As a result, each cyclin-Cdk pair phosphorylates a different set of proteins. Additionally, Cdks can be suppressed by phosphorylation by the kinase Wee1, activated by the phosphatase Cdc25 or inhibited by several inhibitory proteins (CKIs) [14]. The cell cycle control system also depends on transcriptional regulation, where not only cyclins but up to 10% of the whole genome is expressed in a cyclic dependent manner [15].

The first step of the cycle is in late G1 phase (G for gap), which can account for up to 90% of the length of the cycle. G1 phase is a stable state of Cdk-inactivity. This is the preparation phase for mitosis and cytokinesis and if conditions demand it, it can go to G0, which is a steady state. G1 phase ends and S-phase starts with the activation of Cdks via G1/S-cyclins. Now the DNA replication begins at the origin of replication. Helicases unwind the double helix and replication can start. The pre-replicative complex (preRC) is formed by the origin recognition complex (ORC), Cdc6, Ctd1 and a pair of helicases already in G1 phase. It is activated in S-phase via phosphorylation by S-Cdks. At this point, the origin can no longer be used until the preRC forms again during another G1 phase. S-Cdks also phosphorylate the ORC and Cdc6 and thereby inhibit new preRC formation while they are activated. The S-phase (DNA

synthesis phase) lasts 11 to 12 hours in mammals. During this phase, not only are the chromosomes duplicated but the chromatin structure is transferred to the new sister chromatid. At the end of S-phase, sister chromatids are held together by Cohesins, giant ring-like structures. During the following G2 phase the cell expresses genes to prepare mitosis. G1, S and G2 phase together are called interphase. The G2/M transition is the next big step in the cell cycle. Through expression of G2/M cyclins M-Cdks are accumulated. Though phosphorylated, they stay inactive due to inhibition by Wee1. In G2/M transition Wee1 is suppressed and Cdc25 is meanwhile activated and removes the inhibitory restrictions of M-Cdks. A double positive feedback loop occurs – M-Cdks activate their own activator Cdc25 and inhibit their own suppressor Wee1 – which initiates mitosis, a cascade of reactions that takes up to one hour in mammals. Chromosome segregation and nucleus division are separated into five major phases: Prophase, Prometaphase, Metaphase, Anaphase and Telophase. During **Prophase** the chromosomes condensate and the mitotic spindle assembles: a microtubule-based machinery with two distinct poles, the centrosomes. In **Prometaphase** the nucleic envelope is broken down and the chromosomes attach to the spindle at the plus ends of the microtubules. During **Metaphase** the chromosomes are now aligned and the sister chromatids are attached to opposite poles of the spindle by kinetochore microtubules. Now the kinetochore microtubules get smaller and the spindle poles separate, and both pull the chromatids apart from each other. During this process, which is called the **Anaphase,** four motor proteins play an important role: Kinesin-5 has two motor domains and attaches to antiparallel microtubules, pulling them apart; Kinesin-14 attaches to one microtubule and pulls the antiparallel one apart; Kinesin4/10, which are also called chromokinesins, associate with the chromosome arm and push the arm away from the spindle pole while Dynein links the plus ends of the astral microtubules to the actin cytoskeleton and moves towards the minus end, which pulls the poles apart. The metaphase to anaphase transition is controlled via regulated proteolysis. Key regulator is the anaphase promoting complex (APC/C) or cyclosome, which is activated in mid-mitosis via Cdc20 or Cdh1 in early G2 phase [15]. Ubiquitination and, therefore, destruction of the Securins, which protect the protein linkage between sister chromatids, and activation of S- and M-cyclins end the Anaphase. Now **Telophase** starts and the sets of daughter chromatids reach the poles and decompensate. The division of the cytoplasm begins with contraction of the contractile ring. The positioning of the contractile ring during cytokinesis is induced by the positioning of the spindle poles. Now membrane-enclosed organelles are distributed, the ER is separated in two, the Golgi is reorganized, and two new daughter cells are formed. At the end of mitosis APC/C activation leads to overall inactivation of Cdks and destruction of geminin. ORC and Cdc6 are

dephosphorylated and the preRC can form again. SCF, another ubiquitination system, is required to regulate CKIs in late G1 phase and depends on F-box proteins. This complex is responsible for activation of S-Cdks and degradation of G1/S cyclins in early S-phase [7]. Obviously, the cell cycle is the result of tremendous intracellular molecular reorganization. The cellular architecture defined by the cytoskeleton and membranes has to be reshaped and organelles have to be moved to assure functional daughter cells. Altogether an orchestra of cellular trafficking.

1.1.2. Cell-cell interactions: Cell junctions and the extracellular matrix (ECM)

Since a pile of cells does not form an organism by default, these cells must make contacts to each other to communicate in order to build a higher three-dimensional structure. These cell-cell and cell-extracellular matrix (ECM) contacts govern the architecture of the organism, its shape, strength and the arrangement of different cell types through controlling cellular movements within the organism. They are critical for each aspect of organization, function and dynamics of multicellular life [16]. Defects of these result in numerous diseases [17]. We can divide cellular interactions by tissue: in connective tissue and epithelial tissues. **In connective tissue** cells are connected to the ECM rather than to each other. The matrix fulfils most functions of mechanical support. The ECM is made and oriented by the cells within it by three major proteins: glycosaminoglycans (GAG) form chains and hydrated gels and consist of proteins of the collagen family; elastin gives elasticity to the tissue; and fibronectin organizes the matrix and binds integrins. Cell-matrix anchoring junctions control cell proliferation and survival. They recruit intracellular signaling proteins at sites of cell-matrix adhesions and mediate response to mechanical stress. There are two main cell-matrix anchoring junctions: hemidesmosomes, formed by α-6-β-4-integrins and type XVII collagen and actin-linked cell-matrix junctions formed by various integrins. Integrins are transmembrane proteins that can link the ECM to the cell membrane. In the case of hemidesmosomes, ECM ligands are intermediate filaments and the intracellular receptor is plactin and BP230. In case of actin linked cell-matrix junctions, the ECM ligands are actin and adaptors are talin, kindling, vinculin, paxilin and kinases [7]. On the other hand, **in epithelial tissue**, cells are tightly bound to each other rather than to the ECM. Consequentially, the ECM is less pronounced and serves mostly as a thin mat called the basal lamina, which consists mainly of collagen and laminin. There are three main types of cell-cell junctions (1) tight junctions, (2) gap junctions, and (3) anchoring junctions. Tight junctions seal gaps between epithelial cells, while gap junctions allow the passage of small, water soluble molecules. Anchoring junctions link cells together and resist mechanical stress. They must be dynamic and are therefore made by cadherins, which

make hemophilic adhesions. These hemophilic adhesions are relatively weak, and there must be many of them present in order to resist mechanical stress. They play a crucial role in cell sorting mechanisms. There are two kinds of anchoring junctions: desmosomes and adherens junctions. Desmosomes connect intermediate filaments between cells and are made of the noncanonical cadherins desmoglein and desmocollin. Adaptors for the intermediate filaments are plakoglobin and desmoplakin. Adherens junctions connect actin filament of two cells with each other. There are classical cadherins on both cells, and adaptor proteins such as α-, β- and γ-catenins link the cadherins to the actin cytoskeleton. These junctions make mechanotransduction possible via the remodeling of mechanical stress from the cell-layer to the actin cytoskeleton [7].

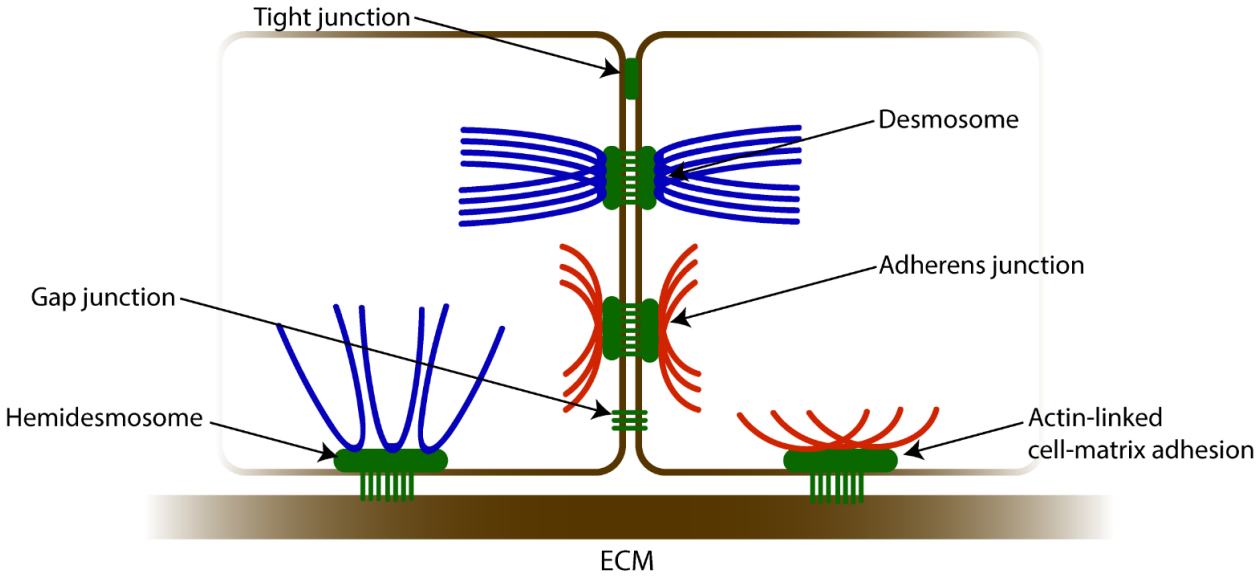


Fig 1. Cell junctions. Schematic presentation of different intercellular junctions and ECM-cell junctions.

1.1.3. Cell specialization: Stem cells, differentiation and tissue renewal

Multiple cells, although connected to each other and surrounded by a supportive matrix, do not always form an organism. For complex organisms to form, numerous different cell types are required to fill various functions, enabling them to form complex tissues. The process of developing specialist cells is called differentiation. The fertilized egg is omnipotent and cells lose potency as they approach terminal differentiation during the cascade of cell divisions. Molecular lifetime plays a crucial part in differentiation during embryonal development. Internal oscillators and external signals are interpreted during cell fate decision. Various gene cascades partially determine the future shape of a cell; one of the most striking examples is the evolutionary conserved *Hox* gene cluster, a genetic machinery to impart specific features to regions along the anterior-posterior axis during development [2], [18]. The overall body structure of every vertebrate has a repetitive, periodic structure. These segments originate from

the mesoderm, which gets segmented in regular somites during embryogenesis. The posterior - the most immature part of the mesoderm slab (presomitic mesoderm) - supplies the required cells: as the cells proliferate, this mesoderm retreats tailward, extending the embryo. In this process it deposits a trail of somites formed from the cells that group together into blocks as they emerge from the anterior end of the presomitic region. The special character of the presomitic region is maintained by a combination of Wnt/Fgf signaling produced at the tail end of the embryo [19]. A conserved gene-expression oscillator acts as a clock to control vertebrate segmentation (*Notch/Delta*) and determines segment size: 30min in *Danio rerio*, 90min in *Gallus gallus* and 120 in *Mus musculus* [20]. This clock is crucial to segmentation and widely conserved between vertebrate. *Hes* genes, which are a key component of the Notch signaling pathway, transcriptionally inhibit Delta and its own expression. *Hes* is expressed until it can inhibit its own transcription, and after degradation it is transcribed again. The timespan of the oscillator and, therefore, the size of one segment depends on the molecular lifetime of *Hes* mRNA and Hes protein. As cells escape the Wnt/Fgf signaling, Notch oscillation arrests in one state or the other. The intercellular Notch signaling pathway keeps this intracellular clock synchronized between neighboring cells.

There are various ways of differentiation, but the most common textbook example starts with a stem cell. Stem cells can divide indefinitely and hold the potential to give rise to both new stem cells and less potential and more differentiated cells. This naturally irreversible process of changing into another more specialist cell is called differentiation. During this process expression patterns change and stemness-associated genes are turned off while cell fate-promoting genes are turned on. This happens numerous times during development and governs self-regenerative processes as well as tissue formation. The size, shape, metabolism and membrane potential of a cell can change during differentiation, while the genetic code stays the same. Reorganization of the cellular framework of course requires intracellular transport and reorganization of organelles, proteins and RNAs. During embryonic development totipotent zygotes and early blastocysts give rise to pluripotent embryonic stem cells, which themselves give rise to multipotent stem cells, oligopotent and unipotent niche stem cells, which are kept for a lifetime and that can give rise to just terminally differentiated cells. Changing the transcriptome of a cell requires transcription factors and more than a decade ago Yamanaka et al. [21] discovered a mixture of transcription factors that are responsible for stemness maintenance. These factors (c-Myc, Oct4, Sox2 and Klf4) seemed not only responsible for keeping stemness but also for inducing stemness in somatic cells and the discovery gave rise to the large field of induced pluripotent stem cell (iPSC) research. During

the last years it became evident that each of these transcription factors can be replaced by a combination of downstream or related transcription factors, chemical reagents and miRNAs [22], [23]. During canonical differentiation cells react to external signals like growth factors with receptors on their cell membrane that lead to phosphorylation cascades within the cell and, finally, activation of one or more cell fate determining transcription factors [24]. The role of signaling during epigenetic control is, without doubt, enormous and the signaling pathways responsible for cell fate decisions are numerous. The most prominent members of this family of factors are: Wnt, Fgf, Bmp, Tgf and Lif, which are all used in cell culture to control in vitro differentiation as well as to culture stem cells. Another way of differentiation is the more intrinsic mechanism of asymmetric cell division in combination with asymmetric distributed maternal determinants. This mechanism is the predominant way of differentiation in organ homeostasis, for example in the liver or gut of mammals. Both genetic systems have a couple of determinants, which all have to be modified during differentiation: pioneering factors like Sox2, Oct4 and Nanog have to be shut down; the polycomb repressive complex (PRC1/2), which is abundant in embryonic stem cells and silences cell fate decision genes, has to be shut down; members of the trithorax group proteins that remodel the chromatin and are responsible for multiple histone modifications have to be reorganized; nucleosomes are repositioned; DNA methyltransferases like Dnmt1 are transiently excluded from the nucleus so that the cells can erase and organize a new DNA methylation pattern; and, last but not least, a number of post-transcriptional control systems are remodeled, for example, the let-7/lin28 system [23].

1.1.4. Cell movement: Internal structure and organization of the mammalian cell

The extraordinary structure of complex organs is the result of the interplay and formation of numerous cells that differ in shape, size, function, membrane potential and metabolism. These differences, results of various differentiation events, are caused by internal structural reorganizations, intracellular transport and localization of proteins and RNAs. As transcription factors and signaling pathways determine the fate of a cell, membranes and the cytoskeleton define boundaries and provide the framework for all cellular functions. And the intracellular organization of a cell determines its function and scope. **Membranes** are crucial for cellular life; they define boundaries between different cells and within the cell as they envelope the nucleus, the endoplasmic reticulum (ER), the Golgi, mitochondria, the lysosome and the centrosome. Ion gradients can be established with semipermeable membranes and can be used to catalyze reactions, to synthesize ATP, to deliver signals, or to transport bigger molecules [25]–[28]. They contain proteins that act as sensors or receptors for external stimuli or factors and therefore enable information transfer between cells or organelles. Their canonical structure

is a 5nm thin bilayer of lipids and proteins that is held together mainly by noncovalent interactions. They are dynamic and fluid structures, two-dimensional liquids with diverse functions. Proteins within the membrane move canonically just within the plane of the membrane and serve as transporters, receptors, links, destabilizers and stabilizers and also the lipid composition of the membrane itself, which is predominantly made of phospholipids, sphingolipids and sterols, changes its functions and characteristics [29].

The **cytoskeleton** organizes the shape, robustness and structure of a cell in space and time and enables the cell to interact with other cells and the environment. Many cells change their shape during lifetime and the basis for all of these changes are molecular rearrangements in the cytoskeleton and intracellular trafficking along the cytoskeleton [30]. The cytoskeleton does mainly consist of three components: (1) actin filaments, (2) microtubules and (3) intermediate filaments. All of these assemble from protein subunits that impart specific physical and dynamic properties and can span hundreds of micrometers. Although the single subunits of all three major components are rather small and diffuse very quickly and additionally are held together by weak noncovalent links, the assembled structures are a solid macromolecular framework throughout the complete cell. Actin and tubulin are both polarized molecules and also form polarized structures, which enable directed intracellular movements that are indispensable to cellular life. Additionally, both provide enzymatic subunits that can catalyze ATP or GTP respectively, which helps to remodel the cytoskeleton more quickly. **Actin filaments** determine the shape of the cell surface and are crucial for locomotion. Due to the polarity of the actin subunits they have a fast growing plus end and a slow growing minus end. They form single helices that underlie the plasma membrane and provide shape and strength to the lipid bilayer, while enabling the cell to contract. They also form static or dynamic cell projections such as lamellipodia or filopodia. Both subunits and the structure of actin filaments are extraordinal conserved between animals. There are several actin binding proteins that vary the properties of the filament; formins and Arp2/3 are responsible for nucleation and therefore assembly of unbranched (Formin) and branched (Arp2/3) networks. Profilin binds actin subunits and speeds the elongation. Tropomyosin stabilizes the filament while various capping proteins, tropomodulin and thymosin, prevent the assembly. Gelsolin severs the filament and only binds to the plus ends. Cofilin binds D-actin and accelerates the disassembly. Fimbrin, α-actinin and filamin crosslink the fibers,thereby strengthening the network as well as ezrin, radixin,moesin proteins (ERM) and spectrin, that link the network to the plasma membrane. Additionally, each actin monomers carries an ATP or ADP molecule and, dependent on the hydrolysis state of said molecule, the filament either grows or is disassembled. There are three

canonical actin homologues in mammalian cells, α-, β- and γ-actin, which vary slightly in their sequence but have very different functions; while α-actin is expressed only in muscle cells, β- and γ-actin are found in almost all non-muscle cells. Actin filaments are often bundled to provide more strength and accession surface for accessory proteins that act as adapters or transporters. Under optimal conditions the filament nucleation is the rate-limiting step in actin formation and once the critical size of the polymer is reached the elongation happens very quickly and monomer availability controls the filament assembly. Intracellular movements of organelles can be powered by actin polymerization and motorproteins can be used to regulate active intracellular transport of molecules along the actin cytoskeleton. Actin-binding motor proteins are members of the Myosin family. Cell migration is vastly governed by actin filament formation at the migrating tip of the cell. The coordination of rapid actin-network formations at these protrusions forms the lamellipodium and sometimes finger-like structures at the leading edge (filopodia) that enable locomotion of the whole cell.

During this process the initial membrane extensions by actin polymerization due to Arp2/3 activity are induced by signaling receptors. Extracellular Egf or Pdgf is detected by Cdc42/Rac transmembrane GTPases and triggers lamellipodia formation at the front edge of the cell. Additionally, lysophosphatic acid (LPA) is detected by Rho transmembrane GTPases,triggering stress fiber formation and contraction at the rear end of the cell. Myosin-II contractions at the rear end of the cell cause organelle and cytosol forward flow within the cell. During retraction of the rear part, cell attachments (focal adhesions) that are rich in signaling receptors are recycled by endocytosis [31], [32]. Each of these processes rely on the correct localization of receptors on the cell surface.

In contrast to the flexible and rather simple actin filaments, tubulin polymers build the stiff and tubular basic framework of the cytoskeleton. **Microtubules** are hollow tubes made of 13 protofilaments that have multiple contacts and a distinct structural and dynamic polarity. They are used to move and anchor cell organelles and molecules at distinct places and build the so-called highways within the cell for long distance transport, for example in axons. They are made of α- and β-tubulin which are frequently found in the cytoplasm and exist in multiple isoforms, making microtubules versatile in their properties. They can form cilia or flagella which are used as motile whips and sensory devices for the cell, with a distinct (9x2)+2 structure and intraflagellar transport of molecules along the tubulin polymers [33], [34]. They can also form bundles through the cell for intracellular transport of various items and emanate from the centrosome in animal cells. Microtubules, though they are complex and large structures, can be rearranged quickly by multiple proteins and do undergo dynamic instability.

β-tubulin binds GTP, which acts as a switch for polymerization. While GTP-β-tubulin tends to polymerize, GDP-β-tubulin rather depolymerizes [35]. Motorproteins make use of the polarity of microtubules and can move either to the plus end (dyneins) or to the minus end (kinesins) of the microtubule. γ-TuRC nucleates the assembly of microtubules and associates with the minus end. TIPs remain associated with the plus ends and can link it to other structures. XMap215 stabilizes the plus ends and Map stabilize the microtubule overall while katanin severs microtubules and kinesin-13 induces disassembly. Stathmin binds tubulin subunits and prevents polymerization, while Map2, tau and plectrin are responsible for filament bundling and cross linking.

Intermediate filaments, in contrast, provide mechanical strength. They are made of small fiber-like proteins that are twisted together using a combination of end-to-end and side-to-side protein contacts. They line the inner face of the nuclear envelope and protect the nucleus. Within the cytosol they twist together and form strong cables that hold epithelial sheets together or help to solidify long structures. All intermediate filaments can provide a platform for motorproteins and can therefore be used for active transport or anchoring of molecules within the cytosol. The order of intermediate filaments within the cell is regulated by providing or depriving attachments within or with other cell structures. The fibers are held together by direct covalent links, and modifications of fiber subunits regulate features of the fiber. They provide a stable and large-scale structure and enable polarity, which often stays for the lifetime of the cell [36].

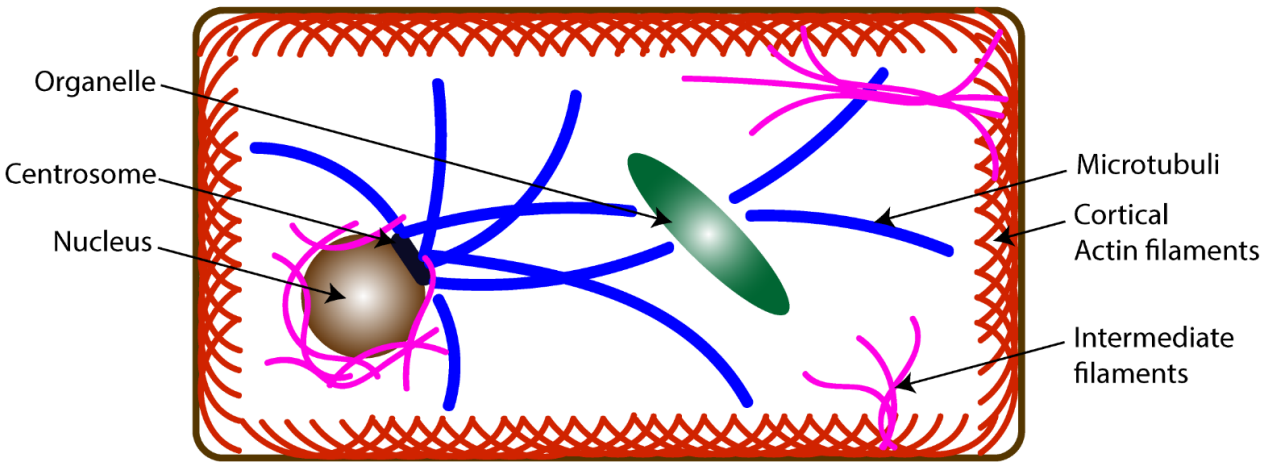


Fig. 2 The cytoskeleton. Schematic presentation of the cytoskeleton.

1.2. Mechanisms and importance of cellular polarity

Cell polarity is the outcome of and reason for multiple cellular processes including junction formation, cytoskeletal organization, organelle positioning, protein and RNA trafficking and

functional specialization of membrane domains. Architecture and shape must be tightly regulated in vivo to fulfil requirements of function and signaling, especially in multicellular life. There are four types of cell polarity: (1) apico-basal cell polarity (ABCP) or epithelial cell polarity; (2) planar cell polarity (PCP) the polarity across the plane of an epithelium; (3) asymmetric cell division (ACD) involved in cell renewal of stem cells and differentiation of daughter cells; and (4) front-rear cell polarity (FRCP) involved in directed cell migration and signaling [10]. Key to all of them is the direct or indirect intracellular transport of molecules and organelles as well as a polar cytoskeleton. The fundamental property of ABCP, FRCP and ACD are antagonistic interactions between two polarity modules, the Scribble homolog module and the Par module [37], [38]. Throughout embryonal development ACD plays a key role in generating different cell types and most cells need to establish functionally distinct domains along an axis of polarity. The interaction of these functional units is essential to multicellular life. The functional units of the cell are proteins, lipids and nucleotides. As in most cellular processes protein interactions are central to the control of cellular polarity and vice versa. Intracellular localization of molecules in various ways is fundamental to all asymmetric distributions and therefore polarization. Moreover, polarized cells need to respond to external stimuli that coordinate polarity at the tissue level. The establishment of asymmetry within the cell relies on intricate networks of molecular interactions between cortically localized proteins. Main complexes of ABCP, FRCP and ACD are the Partitioning defective proteins (PAR) complex (Par3, Par6, aPck) which plays a critical role in apical domain identity together with the Crumbs (Crb) complex (Crb/PalS1/Patj) and the Scribble (Scrib) protein complex (Scrib, Lgl, Dlg), which defines the basolateral domain [39]. The establishment of polarity, however, involves many other proteins including the kinase Par1, the 14-3-3 proteins Par5 and the Coracle group proteins Yurt, Coracle, Neurexin IV, NaK-ATPase as well as at least two types of membrane lipids: phosphoinositides and glycosphyngolipids. The crosstalk to various signaling pathways is crucial since cell polarity relies on extracellular signals and intracellular landmarks that initiate polarization through a signal transduction cascade.

A prominent example of *de novo* polarity formation is the epithelial cyst. Apical markers such as Podocalyxin and Crumbs3 initially localize all around the membrane. Upon cell division they are transcytosed via Rab11 positive recycling endosomes to an apical membrane initiation site (AMIS) that forms around the midbody and is marked by Par3, Cingulin and components of the exocyst. Cingulin is recruited by direct binding to the midbody microtubules and interacts with Rab11 vesicles and Fib5. The fusion of these transcytotic vesicles at the AMIS generates an apical domain which forms a polarized minicyst [40], [41]. A variety of Rab

GTPases are require for trafficking and fusion of these vesicles, where Rab35 is of special importance and binds to the cytoplasmic tail of podocalyxin to capture vesicles around the midbody. Once polarity is established it needs to be propagated and cell junction formation must be coordinated. In most epithelial cells polarity is maintained and transduced by Par proteins (Cdc42, Par6, aPkc, Crumbs, Stardust, Patj, Scribble, Par3, Discs large (Dlg), Lethal 2 giant larvae (Lgl) and Par-1). The maintenance depends on mutual antagonism between apical and lateral factors [42].

aPkc phosphorylates several serines in a basic domain of Lgl that binds membrane phospholipids, introducing negative charges that prevent lipid binding and thereby excluding Lgl from the apical region. Polarity factors can also regulate the localization of other proteins by altering the membrane lipid composition. Phosphatidylinositol 4,5 bisphosphate (PIP2) is enriched apically and phosphatidylinositol 3,4,5 trisphosphate (PIP3) basolateral [43]. Par3 localizes to the apical junction that forms the boundary between these domains and has been proposed to contribute to the differential distribution of PIP2 and PIP3 by recruiting Pten, which catalyzes the conversion of PIP3 to PIP2 and PI3 kinase, which catalyzes the opposite reaction. One landmark of epithelial polarity is the localization of E-cadherin to adherens junctions, which hold epithelial cells together. Par3 is required for cadherin localization in the mammalian epithelium by targeting its exocytosis to the junction domain. The C-terminal, lysine rich part of Par3 recruits the exocyst complex, which is required for the fusion of E-cadherin containing vesicles with the plasma membrane. Polarization has widespread effects on the organization of the cytoskeleton and vice versa. Many components of the initial polarity modules need the cytoskeleton for directed transport and upon activation they reshape the cytoskeleton to fulfil requirements of polarization. Actin is mainly organized locally from sites of cell-substrate or cell-cell adhesions that themselves are controlled by the polarity system. It has, for example, been shown that Cdc42 activates myosin light chain kinase Mrck and thereby induces apical actin-myosin contractility, which plays an important role in the expansion of apical domain and microvilli formation [10], [31], [40], [44].

1.2.1. Cell polarity in neuronal plasticity

Neurons are along the largest and most morphologically complex cells. They are highly polarized, and their function relies on correct intracellular localization of proteins, RNAs and organelles. Those processes enable the distinction between multiple compartments within the cells like the axon, dendrites, synapses, soma and cellular organelles, which all rely on their spatial distinct proteome and transcriptome. Neuronal polarity is fundamental for the correct interpretation and transfer of intrinsic cues and converting them to cytoskeletal rearrangements

or extracellular signaling and vice versa and therefore enables synaptic activity and sculpts neuronal plasticity [45]. The main function of neuronal cells, receiving, processing and propagating signals, is tightly coupled to their morphology. Neuronal polarity is established very early in development and maintained throughout the entire life of a neuron. There are several steps in establishing neuronal polarity and it is most probably the result of a combination of external and internal several cues. The process starts with an **initial, symmetry breaking cue** which is canonically provided by extracellular signaling. This cue gives a spatial orientation and is interpreted by intracellular signaling networks to establish distinct domains. This signaling induces **cytoskeletal reorganizations** and rearrangements that produce the axonal and the dendritic subdomain. Once established, polarity must be maintained throughout the entire life of the neuron [45]. Constant transport of protein and RNA and selective endo- and exocytosis ensure the correct repertoire of factors for each domain. The initial symmetry breaking cue can also be intrinsic, for example in neuronal-epithelial progenitor cells, which are already bipolar cells and give rise to ganglion cells. Mechanosensory neurons can inherit their polarity from the site of cell division of their progenitor cell through an accumulation of Rho1 and AuroraA at the site of cytokinesis, followed by an enrichment of PtdIns (4,5) and P2. The subsequent clustering of D, E-cadherins and Bazooka/Par-3 determine the first site of protrusion outgrowth. Two examples of intrinsic polarity regulation are cortical and hippocampal neuronal cultures [46], [47]. As cells are dissociated before differentiation, they lack the higher organization of the tissue. They start to grow immature and undetermined protrusions into many directions. The first morphological manifestation of polarity occurs in stage 3 when one neurite, which will give rise to the axon, begins to grow faster and extends longer. Upon this event, all other protrusions will become dendrites. Though, the lack of external symmetry breaking factors in neuronal development is rather unlikely since correct and directed neuronal wiring is critical for neuronal network formation. In most situations there are external factors that provide a compass for polarization, which is usually a gradient of a stimulus and the correct interpretation of that stimulus directs polarization. Cortical pyramidal neurons, for example, are ‘born’ in the ventricular zone. They migrate apically along the radial glial processes and at the subventricular zone they start to extend several neurites. At this step they are called multipolar. Polarity is initially manifested during migration when one minor neurite grows rapidly and becomes the leading process, upon which all the other outgrowths become trailing processes. Those minor neurites retract to give rise to a bipolar cell in the intermediate zone. Importantly, the leading and trailing process eventually gives rise to the apical dendrite and axon, respectively. Therefore, the polarity is generated via directed neuron

migration [48]. Tgf-β, which emanates from the ventricular zone, is the main diffusible cue that induces cortical neuron polarity and is also in vitro sufficient to stimulate axon specification and growth. Other prominent diffusible cues are Bdnf and Nt3 which are important for axon induction and multipolar to bipolar transition, where Bdnf acts in an autocrine or paracrine way to self-amplify. In about one third of all cases the leading edge (future dendrite) appearance precedes that of the trailing edge, indicating that polarity can be manifested prior to axon specification by the aforementioned signals [46]. N-cadherin, which is expressed in a polarized manner in newborn neurons at the SVZ, promotes first neuronal outgrowth in vitro and is required for multipolar to bipolar transition. Since all of these factors are also involved in migration, the interpretation of cause and effect knockdown or knockout studies is problematic. Axon-axon contacts are also important for early axonal development. *Tag1* is expressed by mature neurons.Multipolar cells grow their axons along the Tag1 tracks of existing neurons. Substrates of the ECM, such as laminin, tend also to promote axon formation. The laminin receptor Itgb1 is required for axonal outgrowth and was suggested to regulate polarity via Lkb1 pathway and microtubules. Additionally, at least four different intracellular signaling pathways are crucial for the establishment of neuronal polarity: Tgf-β, Bdnf, Wnt and Rac. Several positive and negative feedback loops control external and intrinsic signals. For example, the local enrichments of the Wnt receptor Lin-17 via the Wnt ligand Lin-44, upon which does response to the Wnt signal and leads to signal amplification. Another example of signal amplification is Bdnf-triggered Bdnf secretion, which leads to a cascade of phosphorylations within the cell. Also, negative regulation of axonal fate in future dendrites is controlled in various ways. While cAMP promotes axonal growth, cGMP promotes dendritic fate [49]. Calcium waves from the primary axon act inhibitory on other outgrowths [50], [51]. The anterograde transport of a polarity regulator, coupled to its retrograde diffusion, leads to its concentration in the longest process [52]. *RhoA* inhibition in vitro leads to suppressed neurite outgrowth and inhibition of the downstream Rho kinase *Rock* leads to multiple axon formation [53], [54].

In mammalian neurons, the small GTPases Cdc42 and Rac1 link the Par complex to actin. Activated Cdc42 is part of the Par-3/6 complex in neurons. Indeed, Cdc42 mutants have shorter axons and in cell culture *Cdc42* KO cells do not have any axons, show aberrant actin dynamics and excessive phosphorylation of the actin regulator cofilin, which is also required for polarity induction [55], [56]. Actin regulation via Par-3/6 complex, *Rac* activation and recruitment of the GEFs Tiam1/Stef together lead to over-activation and inhibition of Par3/6 and impair

neuronal migration and multipolar to bipolar transition. *Rac1* mutants also show severe defect in axon formation [57].

Microtubule formation is also regulated by polarity signaling. The Lkb1 pathway induces phosphorylation of Map and Tau via SAD kinases and thereby reduces Tau binding to microtubules. Phosphorylation of Stathmin/Opt18 inhibits its microtubule destabilizing activity and leads to various cytoskeletal rearrangements. As stated, cytoskeletal regulators are main downstream targets of polarity signaling. Microtubules are extended in general and sorted during neuronal polarization. Axon microtubules are mainly 'plus end out' whereas they are not sorted in dendrites. This is usually used in immunofluorescence to label axons and dendrites: Map2 is typically used to label the somatodendritic part and dephosphorylated Tau to label the axon [58]. Neuronal microtubules have many protein modifications, are long lived and are much more stable. Their modifications affect mainly the kinesin dynamics. In young neurons a massive nucleation of microtubules in the newly formed axons occurs, leading to many and stable microtubules that can be used for faster and more reliable molecular transport to the axon and therefore faster growth of the protrusion [59]. Neuronal polarity is closely linked to microtubule polarity. It is not completely known how microtubule polarity is established, however there are several known regulators of microtubule polarity: (1) Trim46 bundles parallel microtubules and its knockout leads to Map2/Tau diffusion and 25% of microtubules growing in the wrong direction [60]; (2) Dynein may be required to establish uniform microtubule polarity by transporting plus-end-out microtubules into the axon and removing minus-end-out microtubules [61]; (3) in *C. elegans* Crimp2 homolog Unc33 is required to establish microtubule polarity. Unc33 binds tubulin dimers and is transported by Kif5 [62].

The **actin cytoskeleton** is reorganized too. Axons of mature neurons contain periodic sub-membrane actin rings that are organized by beta-spectrin and adductin, whereas dendrites contain way fewer rings [63]. Another actin structure of the axon is a mesh-work at the axon initial segment (AIS) that is thought to regulate transport in and out of the AIS [64]. Additionally, actin dynamics play a crucial role in the early stages of polarity establishment: Axonal growth cones show more dynamic actin than the minor neurites. Cofilin is enriched in the axon growth cone and destabilizes actin filaments in order to allow microtubule growth into the future neurite [65]. Also, Cdc42 destabilizes actin and therefore promotes microtubule growth into the axon. Rac1, Cdc42 and actin waves from the soma towards the growth cone result in higher concentration in the axon than in the dendrite and rhythmic growth at the growth cone [31].

Despite those molecular mechanisms shaping neurons and therefore the synaptic network, it was shown that local protein synthesis is crucial for memory formation and synaptic plasticity. Not only does the synaptic potentiation by Bdnf require local translation [66] but also long term facilitation in aplysia [67], long-term depression facilitated by metabotropic glutamate receptor activation [68], late-phase long term potentiation (LTP) [69], dopamine-induced plasticity [70] and homeostatic plasticity induced by a blockade of spontaneous neurotransmitter release [71]–[73]. However, in most cases the specific proteins are not known.

1.2.2. In the wrong place at the wrong time: Cellular polarity in development, disease and cancer

Cell polarity and its fundamental processes, the intracellular localization of proteins and mRNAs, and the organization of internal and external structure are crucial for multicellular life. Disturbances in polarity govern multiple diseases and play an important role in several forms of cancer. Many perturbations in cell polarity genes lead to cancer, whether they are mutations, transcriptional or post-transcriptional deregulations. In fact, the cell polarity modules Scribble (Scrib), Discs-large (Dlg) and lethal-2 giant larvae (Lgl) have a tumor-suppressive role in mammalian epithelial cancer [37], [74]. They play a scaffolding role in various pathways and are involved in tissue growth, differentiation and cell migration. Also, the Par module and Crb are deregulated in many forms of cancer [75]. The tumor suppressive roles of Dlg, Scrib and Lgl were first discovered in Drosophila melanogaster [76] and result in excessive cell proliferation and formation of neoplastic tumors. There are four mammal orthologues of Dlg (hDlg, PSD-93, NE-Dlg and Psd95), two of Lgl (Lgl1 and Lgl2) and only one of Scrib, which makes the functional analysis of Scrib in mammalian systems more amenable. In mice it suppresses epithelial-mesenchymal transition (EMT) and the knockout is embryonic lethal [77]. The tissue specific KO of Scrib in mice leads to hyperproliferation and cell morphology changes, leading to tumorigenesis in prostate, breast, lung and skin epithelial tissues. KO studies of Dlg1-4 or Lgl1&2 are more difficult, yet result in defects of PCP and therefore failures in neuronal tube closure and sheet migration, which are embryonic lethal [38]. It is not surprising that disturbances in those main polarity regulators are fatal to mammals, however defects in downstream localization of specific mRNAs and proteins are the cause of many genetic diseases. **Myotropic dystrophy**, for example, is caused by an expanded CTG or CCTG microsatellite repeat in the 3’ UTR of the dystrophia myotonica protein kinase gene (*Dmpk*) or the first intron of CCHC zinc finger nucleic acid binding protein (Cnbp). Both mutations act as scaffolds and lead to a recruitment of Mbnl into nuclear foci. Mbnls are regulators of alternative splicing and are involved in fetal-to-adult isoform transition. Mbnl1

and Mbnl2 are known to act on mRNA stability and localization and the membrane-bound local translation in numerous cases [78]–[81]. Loss of function of those regulators leads to various mis-splicing events and causes multiple phenotypes of myotonic dystrophy by deregulation of intracellular localization of: (1) chloride channel 1 (Clcn1) causing myotonia [82]; (2) insulin receptor (IR) causing insulin resistance [83]; (3) bridging integrator 1 (Bin1) causing muscle weakness [84]; (4) calcium channel (Ca(V)1.1) causing muscle weakness [85]; (5) sodium channel 5a (Scn5A) causing cardiac arrhythmia [86] and cardiac troponin T (CTnT), also causing cardiac arrhythmia [87].

The dysregulation of mRNP assembly and localization changes the fate of multiple mRNAs, since it governs all steps of posttranscriptional regulation. Hyper- or hypo-assembly of mRNPs can lead to **spinal muscular atrophy** (SMA) or **amyotrophic lateral sclerosis** (ALS) [88]. Alhough the formation and composition of stress granules is well characterized, the regulation and transport of mRNPs is not well understood. Defects in the survival of motor neuron (SMN) complex, which acts as an assembly machine for multiple RNPs [89], leads to SMA. As the SMN complex interacts with multiple RBPs and is part of RNA transport granules in axons, it is likely that SMN has a noncanonical role in mRNA transport or metabolism. Mutations in key components of the SMN complex result in defects of axonal mRNA transport and local translation for example of Gap-43 and Cpg15 [90].

The **fragile-X-syndrome** (FXS) might be the most studied example of a localization defect leading to a genetic disease, since it is the most frequent mental retardation after Down syndrome. FXS is characterized by increased synaptic protein synthesis and stimulus-insensitive synaptic protein synthesis, causing defects in synaptic plasticity. It results from deregulations in transcription or localization of Fmrp, which itself is an RNA-binding protein responsible for spatial, activity-dependent protein synthesis [91]. Fmrp regulates the activity of P13K, a component of the mTor pathway, which is involved in transmitting external stimuli into intracellular protein synthesis in the synapse. Apart from general dendritic protein synthesis, the mTor pathway is responsible for the regulated translation of CaMKIIa.

Autism, which itself is a mixture of mental disabilities, can be caused by mutations in *Rbfox1,* which has a nuclear and a cytoplasmic isoform (paralogues are NeuN and Rbm9). Both proteins are master regulators of autism-associated genes and have well characterized functions in alternative splicing. Additionally, the cytoplasmic isoform of Rbfox1 binds target mRNAs and increases their stability and might have an antagonizing function to miRNAs. It is responsible for the expression of multiple synapse-associated genes [92].

In general one can say that granule formation, mRNA localization and spatial translation are crucial for neuronal functions, plasticity and survival [93]–[97]. The regions of proteins involved in granule formation are often low-complexity regions and mutations in those regions are linked to several and severe mental and neurodegenerative diseases [98].

1.3. The "triple T" destiny of an mRNA: Transcription, Transport, Translation

The fate of an RNA is controlled in various ways. Its origin, the process of transcription, is a complex and highly regulated process. Its destiny is regulated by multiple cis- and trans-acting factors and its functions vary from delivering 'just' the message for protein production, helping enzymes in target recognition or having itself a catalytic activity. All RNAs are polymerized by RNA polymerases either in the nucleus or in mitochondria. Depending on the type of cell, the developmental stage of the cell, extracellular cues and of course the message itself, mRNAs are transcribed, transported and translated through various molecular mechanisms, tightly regulated by internal sequences and modifications.

1.3.1. Transcription: The birth of a new RNA

There are three different RNA polymerases: RNA polymerase 1-3 (RNA-PolI, RNA-PolII & RNA-PolIII), which all recognize different promotor sequences. RNA-PolI, localized in the nucleoli and transcribes most ribosomal RNAs (rRNAs). Its promotor is a highly species-specific bipartite promotor and spans -31 to +6 region of the gene and therefore overlaps with the gene itself. Efficient transcription also requires an upstream promotor element [99].

RNA-PolII is responsible for most mRNA transcription events. It is localized in the nucleoplasm and can recognize several complex and diverse promotor sequences. The most common elements of promotor regions for RNA-PolII are: (1) CpG islands, GC-rich regions upstream from the actual promotor regions, which are common for housekeeping genes; (2) TATA-box elements, AT-rich sequences located 25-30nt upstream from the transcription start site and bound by the TATA binding protein (Tbp); (3) downstream promotor elements (DPE); (4) transcription factor II B (TfIIB), recognition element (BRE); (5) motif ten elements (MTE); and (6) the initiator element (Inr), which includes the four first initiating nucleotides and exists in about half of all genes recognized by RNA-PolII. In addition to promotors there are enhancers, trans-acting transcriptional activators that can have variable positions and orientations and are recognized by transcription factors that stimulate RNA-PolII binding to the promotor [100].

RNA-PolIII is also located in the nucleoplasm and transcribes precursors of the 5S rRNA and tRNAs as well as a variety of other nuclear and cytosolic noncoding RNAs [101].

The process of RNA-PolII transcription is well studied and starts with the formation of the preinitiation complex (PIC). Tbp binds to the TATA box element and kinks and partially unwinds the DNA so that transcription factor II D (TFIID) can bind. TfIIA, TfIIB and the tata binding protein association factor (Taf) interact with TBP and RNA-PolII, and other initiation factors join. The composition of canonical and noncanonical transcription factors in the PIC determines the regulatory ability, but at least Tbp, TfIIB, TfIIE, TFIIF and TfIIH must bind to initiate transcription. Canonical transcription factors are: (1) TfIIA, which stabilizes Tbp and Taf binding; (2) TfIIB, which also stabilizes Tbp and recruits RNA-PolII; (3) TfIID, TBP and TAF, which recognize the TATA-box element, recruit TFIIA and B and have regulatory functions; (4) TFIIE, which recruits TFIIH and stimulates its helicase activity, enhancing promotor melting; (5) TFIIF, which facilitates promotor targeting and stimulates product elongation; and (6) TfIIH, which is an ATP-dependent helicase and which melts the promotor so that RNA-PolII can initiate. Once the PIC is formed, RNA-PolII requires different elongation factors to bind the phosphorylated CTD of Rbp1 and accelerate its transcription. Unlike prokaryotic transcription, which has a determined termination sequence, eukaryotes lack the determined termination. Premature mRNAs further undergo co-transcriptional splicing events and posttranscriptional modifications, such as capping and A-tailing, which determine the further future of the transcript [7].

1.3.1.1. Splicing and alternative splicing

The most striking difference between prokaryotes and eukaryotes concerning the genome is the structure of genes with introns and exons in eukaryotes compared to uninterrupted genes in prokaryotes. In eukaryotes, pre-mRNAs are variable in length and much larger than expected from the protein size (2-20kb). They are processed by cutting out noncoding regions, or introns, and ligating the neighboring coding regions, orexons. This process is called splicing. In humans, the number of introns varies from zero to 364 within a single gene, with intron length varying from 65nt to 800,000nt. Exons vary in length between 150nt to 17,000nt [7]. The production of functional mRNAs starts with the transcription of the full gene. Capping occurs soon after initiation and splicing starts already during the process of elongation. The mature RNA emerges after poly-adenylation and splicing and is transported to the cytosol. Splicing is a two-stage reaction in which, at first, a lariat structure of the intron is formed, which is eliminated in the second step when the exon junction is ligated. First a 2',5'-phosphodiester bond forms between an intron adenosine residue and the introns 5' terminal phosphate. The 5'

exon is thereby released and the intron assumes a lariat structure, in which the adenosine at the lariat branch point is typically located within a conserved sequence 20 to 50 residues upstream of the 3' splice site. In the second step the free 3'OH group of the 5' exon displaces the 3' end of the intron, forming the phosphodiester bond with the 5'-terminal phosphate of the 3'exon yielding the splice product. The intron is thereby eliminated from the structure, linearized and rapidly degraded. The invariant GU at the intron's 5' boundary and the invariant AG at its 3' boundary are necessary and sufficient for a splice junction. Splicing occurs without free energy input. Since splicing happens during transcription, skipping of a splice junction and, therefore, skipping of an exon is rare. The process is mediated by snRNPs in the spliceosome. The spliceosome consists of nuclear ribonucleoproteins, specifically 5 RNAs and 150 polypeptides, and is comparable in size to the ribosome. All spliceosomal RNPs (Ux-RNPs) have the same core-proteins, which are called Sm proteins, that bind the Sm motif. U1-snRNP is outstanding and partially complementary to the consensus sequence of the 5' splice junction and therefore recognizes the splice junction. Splicing offers evolutionary advantages since it enables proteins built of so-called "functional blocks" to be recombined [102], [103]. Alternative splicing can form multiple proteins from one or more genes. SR proteins (Serine – Arginine proteins) carry an RNA recognition motif (RRM) and bind specific exonic splice enhancers (ESEs) and recruit the splicing machinery in both canonical and alternative splicing processes. Contradictory hnRNPs bind to intronic or exonic splicing silencers (ESSs /ISSs) and block the binding of the splicing machinery. The combination of ESEs, ESSs, ISSs, SR-proteins and hnRNPs determines the predominant splicing variant. There are various different events combined in the term of "alternative splicing" and more than 95% of all mature human RNAs are the outcome of at least one of them [104]. Intron retention is, as the name suggests, the retention of an intronic sequence in the mature mRNA, whereas exon skipping describes the exclusion of an exon in the mature RNA. Circular splicing results in circular RNAs (circRNAs) and is the ligation of a splice donor site to an upstream acceptor site. Cryptic or non-canonical splice sites can be used just under specific conditions and vary the mature RNA outcome, whereas in trans-splicing a splice junction of one transcript is ligated to another transcript. All of these processes are utilized by eukaryotes to modify, regulate and coordinate their transcriptome [105], [106]. Failures in splicing lead to noncanonical mRNAs and therefore nonfunctional or malfunctional genes. Correct splicing is crucial especially in the context of RNA transport, since the exclusion or inclusion of RBP binding sites alters RNP formation and therefore the intracellular fate of the mRNA. It has been shown that alternatively spliced transcripts often show different localization patterns and that this process is used by the cell to control the spatial

transcriptome [79], [98], [107], thus post-transcriptional modifications in general are widely used to control gene expression in neurons. The impact of alternative splicing on differential localization and local translation was part of various studies [108].

Shigeoka et al. found that the axonal translatome showed more diversity of mRNA isoforms than the somatic translatome [109]. They used their dataset of axonal ribosome associated mRNAs to test in silico 164 alternative events that produce two isoforms both in the axonal and in the retinal compartment. Though they could not detect a general trend in translation efficiency towards the longer isoform as it was postulated by Taliaferro et al. [110], they observed a clear bias of the PSI (percentage spliced in) value calculated by MISO [111] towards the two extremes, indicating that only one of the two isoforms is preferentially translated in the axon. Unfortunately, they do lack information about the general abundance of mRNAs in the compartments and cannot distinguish between isoform-specific localization or local translational regulation. Additionally, they found several back-splicing events resulting in circular mRNA isoforms in the axonal compartment pinpointing towards the translation of circRNAs in the axon. Using a FRAP (fluorescent recovery after Photobleaching assay) with a diffusion-limited fluorescent reporter they identified two alternative exons of the genes Acot7 and Stx3 that seem responsible for axon-specific translation of the reporter.

In 2018 Tushev et al. published an extensive alternative 3' UTR analysis of neuronal compartments. Using rat hippocampal microdissection to separate the neuropil layers from the somatic layers and employing a sequencing technique enriching for 3' UTRs, they identified a huge diversity in mRNA transcripts, showing specific enrichment for a particular 3' UTR isoform [112]. In contrast to Shigeoka et al., they found that 3' UTR isoforms of neuropil localized transcripts are significantly longer and do inherit more and duplicated regulatory elements. Finally, they state that the neuropil-enriched 3' UTRs isoforms have significantly longer half lives.

1.3.1.2. Capping and poly-adenylation

Eukaryotic mRNAs carry 5' caps, m7G residues joined to the 5' end of the transcript via 5'-5' triphosphate bridge formation. These are added to the transcript co-transcriptionally when the transcript is about 30nt long and identifies the translation start site. Capping involves several enzymatic reactions: (1) the removal of the leading phosphate group by an RNA triphosphatase; (2) the guanylylation of the mRNA by the capping enzyme; (3) the methylation of the guanine by guanine-7-methyltransferase; and (4) sometimes also be $O^{2'}$-methylation at the first or second nucleotide. It protects the transcript from 5' exonuclease activity [113].

On the other side the transcript is 3' A-tailed. Since the 3' ends of transcripts are imprecise due to the nature of termination, A-tails are added to uniform and protect the 3' ends. Despite the imprecise 3'end of the premature transcript, they usually start at specific positions and are added in a two-way reaction. The transcript is cleaved 15 to 2nt after the AAUAAA sequence by cleavage factors I and II (CfI and CfII). Subsequently, the poly-A-tail is added by poly(a)polymerase (Pap), which acts independent of templates and is activated by cleavage and polyadenylation specific factor (Cpsf). Once the adenosine-chain has been elongated to about 10nt the AAUAAA sequence is no longer needed. Cpsf binds to the phosphorylated CTD of RNA-PolII and therefore couples polyadenylation to transcription. The length of the A-tail is determined by the binding of multiple copies of the poly-a-tail binding protein II (PabpII) [114]. The A-tail is a protective end for 3' exonucleases and shortens with the lifetime of an mRNA. It is bound by the poly-A binding protein (Pabp), which protects the mRNA and is also involved in translational regulation [115]. Pabp also regulates intracellular localization of the messenger RNA. Although Pabps predominantly diffuse and shuffle freely between the nucleus and the cytoplasm, the sub-cellular distribution can alter dramatically upon viral infection or other forms of cellular stress [116]. Very short life mRNAs do not have A-tails, for example histone mRNAs with an average lifetime of under 30min.

1.3.2. Importance and mechanisms of RNA localization

The spatial environment of a functional unit within the cell alters its functional potential. Therefore, many proteins and RNAs are not evenly distributed within the cytoplasm but are localized to specific cellular compartments to fulfil the required functions. There are several advantages of RNA localization over protein localization: (1) there are fewer mis-localized proteins if the translation happens on the spot; (2) one RNA can give rise to several proteins and therefore the localization of an mRNA is energetically more efficient for the cell; (3) the cell can react more quickly to external stimuli, which is of special importance for huge cells like neurons; and (4) spatially translated proteins do carry compartment-specific modifications and can therefore differ from unlocalized translated protein products. The regulation of transport and spatial translation are critical to all aspects of eukaryotic life. There are several mechanisms known for the localization of transcripts, which can be divided by directionality. There are many forms of direct and active transport described in literature either for single mRNAs via adaptor RNA binding proteins (RBPs) and motor proteins, as batches of transcripts via transport granules, multi-RNA and multi-protein aggregates that can also be transported via motor proteins along the cytoskeleton [96], [117]–[119] or via anchoring to cargo vesicles

or organelles that do travel along the cytoskeleton [26], [28], [120]. In contrast to these mechanisms, intracellular inhomogeneity of transcripts can also occur indirect via pre-localization of effectors. In detail, these effectors mediate mechanisms such as spatial degradation, local translational regulation, or local anchoring. All these modes of localization require the correct identification of a mRNA and interpretation of the internal localization signal. Additionally, active transport requires a molecular motor, which is added to the mRNA, granule or organelle and which can direct the cargo to its destination. Some mechanisms require translational downregulation during the transport and spatial de-repression while others can occur co-translationally [34], [117], [121].

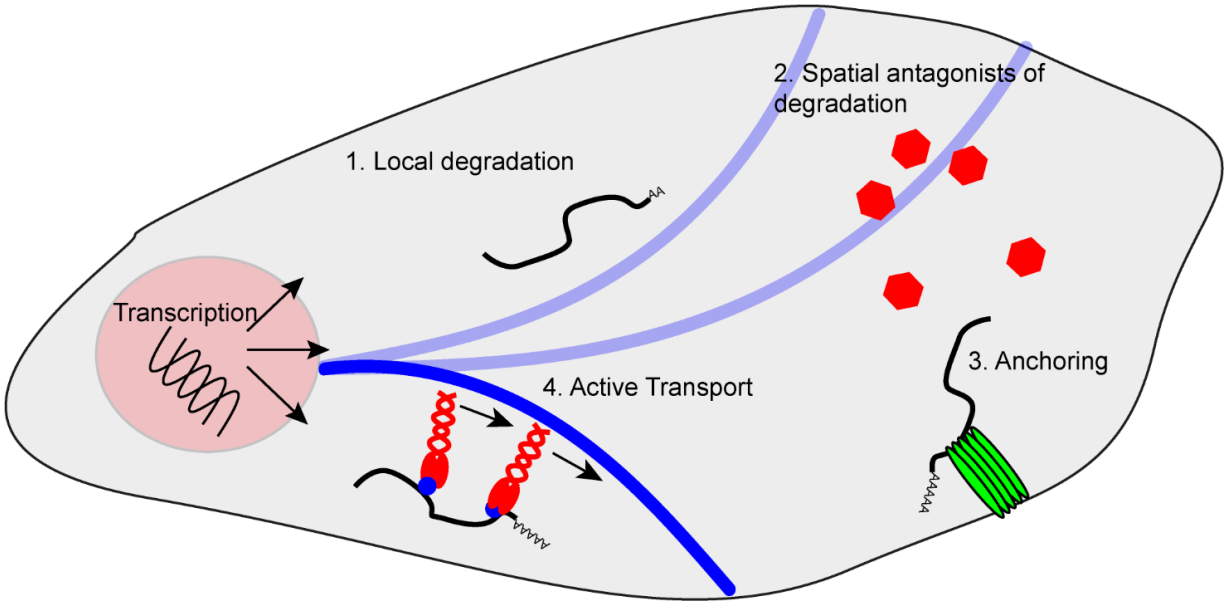


Fig. 3 Methods of mRNA localization. Spatial mRNA enrichment can be achieved via local degradation, local mRNA protection, diffusion and spatial anchoring, active transport and all combinations of those phenomena.

1.3.2.1. Proteins that bind and manipulate RNAs in versatile ways (RBPs)

Although ribonucleotides can have a catalytic function themselves, they do require proteins for most of their functions, which are the predominant active molecules within the cell. RNA binding proteins (RBPs) are, in general, all proteins that bind to RNA and can therefore potentially change the fate or function of the RNA. They generally need an RNAbinding domain or adapter proteins to contact the RNA. The predominant form of an mRNA in the cell is never the unbound state, as it might be in some biochemical assays, though in nature mRNAs act more like scaffolds and are packed with a dynamic coating of various proteins. In most cases it is the protein that changes properties of the RNA, but this can also happen the other way around. Very prominent examples are sgRNAs in the CRISPR/Cas9 system or miRNAs that act as adapters or guides and help the protein to recognize its targets. RBPs can generally

be grouped by the following functions: translational regulators (repressors, de-repressors and initiators); nucleotide metabolism regulators (exo- and endonucleases, ligases, polymerases and members of the splicing machinery); RNA modifiers (for example RNA methyltransferases, A-tailing enzymes or the capping machinery); structural modifiers (for example helicases); and RBPs involved in localization (adaptor proteins that link RNAs to motor proteins, for example). RBPs can also be grouped by their way of binding. There are numerous known canonical binding motifs, such asthe RRM, KH, RGG, C2H2/CCCH (zinc finger), dsRBP, PUF, PAZ and DEAD-box proteins and, on the other hand, non-canonical RNA binding motifs like DZF, PDZ, NDR, HSPA1A [122] or unstructured or intrinsic disordered regions (IDRs). They can bind either the bases of the ribonucleotides, which makes them easier to detect via crosslinking methods, or the ribose backbone, which makes them barely detectable by crosslinking methods. A third way of classifying RBPs is by the RNP that is formed. In cell biology it is widely accepted to speak about cajal bodies [123], paraspeckles [124], processing bodies [125], stress granules [126] and transport granules [127]. The mRNA and RBP composition of these RNPs is, in many cases, redundant. Also, the term granule is misleading, since it normally refers to solid phase objects, while RNPs can occur in versatile states and are very often liquid phase separations. The binding of RBPs to RNAs can occur either in a sequence-specific manner, by relying on a structural motif, a combination of both, or in a completely non-specific manner. In mice, about 1400 RBPs have been identified so far and experiments have shown that the RNA protein interactome is highly dynamic and can change dramatically due to stimuli [122]. A lot of monogenic diseases are the result of mutations in either RNA binding domains or in RBP-recognition elements that impair both the correct interaction of a specific RNA and its RBP [128]–[130]. A prominent example of these mechanisms is the survival of motor neuron protein (Smn), which is part of a multi-protein complex regulating spliceosomal small nuclear ribonucleoproteins (snRNPs). Low levels ordysfunctional Smn result in spinal muscular atrophy, a motor neuron disease. Smn is known to associate with various RBPs, although the nature of this interaction is unknown. Donlin-Asp and colleagues found in 2017 that Smn proteins promote the interaction of Imp1 proteins and the 3’ UTR zip code of β-actin, leading to the assembly of the trafficking mRNP that associates with the cytoskeleton [131].

1.3.2.2. Zip codes: RNA sequences that regulate mRNA localization

The **sequence specific binding** of RBPs to mRNA, which leads to localization, is mediated by zip codes, that is, RNA elements that mediate localization in cis. This term was first introduced in the early 90’s by Robert Singer and colleagues [132] and refers to a sequence of the RNA

that regulates its localization. Modifications, either in the trans acting RBP or the cis acting RNA element, can lead to defects of localization. Localization elements (LE) or zip codes are typically encoded in the 3'UTR of the mRNA, although they can also be found within the ORF. They can vary in size from a few nucleotides to over 1kb and they can be recognized by various RBPs or just one [118]. One of the best-studied examples is the **β-actin zip code**. It is a 54nt long sequence in the 3' UTR of the mature mRNA. It was identified by narrowing down the 3'UTR of β-actin fused to a reporter RNA [133]. The motif is bound by two RBPs with different binding mechanisms. The core motif is a 28nt bipartite localization element forming a stem loop structure that is recognized by ZBP1 (zip code binding protein 1). Additionally, HuD binds an A-rich part of this motif sequence specific. β-actin requires the presence of the zip code in its 3'UTR to be targeted to the leading edge of migrating fibroblasts, dendritic filopodia and axonal growth cones. Zbp1 is the key factor that binds the *β-actin* mRNA in the nucleus and is involved in its localization and translational repression. The hnRNP homolog Ksrp binds beta actin mRNA subsequently after maturation and facilitates nuclear Zbp1 association. HuD joins the complex once the mRNA leaves the nucleus [134].

Other well-studied examples are the mRNA of **myelin,** which contains an 11nt long element called A2RE in its 3'UTR that is recognized by hnRNP and facilitates transport into oligodendrocytes, or **bicoid** mRNA in drosophila melanogaster, which is recognized by egalitherian, an RNA adaptor protein involved in the dynein mediated transport. ***Ash1*** in yeast has four distinct zip codes (E1, E2, E2B and E3) that do not share homology but mediate localization through the same RBP, She2p, which enables myosin-based transport [135].

No clear pattern of zip codes is known so far, which is unsurprising, since there are numerous ways of localization mediated by different RBPs. First protein coatings of mRNAs are implemented already by pre-mRNA processing in the nucleus and, once in the cytosol, the ribonucleotide particle can gain or lose more RBPs. In terms of transport, these factors may define stability and translatability and mediate the contact to motor proteins or anchoring proteins. mRNAs are often transported in large and diverse multiprotein complexes. These complexes might also contain miRNAs and lncRNAs [93]. Defining which RBP binds to which zip code and, therefore, which RBP binds to which mRNA, holds critical information about how mRNA trafficking and localization is regulated in different cellular environments. The identification of zip codes *de novo* involves accurate knowledge about localization of mRNAs to subcellular compartments and a precise reporter system. The predominant RNA movement in the nucleus is diffusion. Active transport of mRNAs occurs for the first time when they leave

the nucleus. This process takes about 0.2s but the mRNA remains bound to the nuclear membrane for seconds, for reasons of quality control and RNP formation. Following that, the fate of the mRNA is either free diffusion through the cytoplasm or anchoring to a transport system. In all studied cases of RNA transport it is a combination of multiple mechanisms; the average outcome of all processes results in a inhomogeneous distribution of the mRNA[119]. Single RNA tracking of lacZ mRNA in mammalian cells, for example, revealed that 50% of the mRNA is freely diffusing, while 25% remains confined by microtubule base domains, 20% stays associated with the cytoskeleton and just 5% moves along the cytoskeleton [136]. Oscar mRNA in *drosophila melanogaster* during embryogenesis shows a similar pattern. The localization is regulated by several trans acting factors such as staufen and hnRNP A/B, which mediate active transport via kinesins that move the RNP along the microtubule cytoskeleton [34], [137]. However, microscopy revealed that most mRNAs are diffusing and only 13% are actively transported – and 7% are also transported to the wrong pole due to misdirected microtubules [138], [139]. Therefore, the directionality and organization of the cytoskeleton defines the accuracy and efficiency of directed mRNA transport, which is the case for the clear majority of actively transported RNAs. It is speculated that the preferred way of localization in most cells is free diffusion through the cytoplasm, coupled to local entrapment via anchoring proteins rather than energy-consuming active transport [30].

There are several examples known for both phenomena in the *drosophila* oocyte. *Oskar* and *Nanos* are transcribed by nurse cells in the *drosophila* embryo and transported to the poles of the oocyte. While *nanos* mRNA is freely diffusing and anchored at the poles of the oocyte [140], *oskar* mRNA is continuously transported to the posterior pole via kinesin-1 [141]. Though the active localization is just slightly biased towards the anterior end, the overall distribution is very distinct due to additional anchoring of *oskar* upon arrival at the destination [142]. The different mechanisms in this case might indicate the importance of contemporary transport. *Bicoid* mRNA that is localized to the anterior pole and induces head formation exhibits a combination of localization events in the *drosophila* embryo. It is randomly transported via dyneins in particles of constant size but varying RNA content. Since the microtubule cytoskeleton at the anterior part of the oocyte is not directed, it cannot be localized just through transport. It has been shown that the maternal coating of the mRNA from the nurse cells is crucial for its correct anchoring at the anterior pole via the pseudonuclease exuperantia (Exu), which mediates spatial P-body sequestration [143].

There are three known distinct ways of local mRNA anchoring: (1) active transport via microtubules and anchoring via actin binding, which requires an adaptor switch at the destination[34], [144]; (2) turning the motor off at the destination – switching motor activity of the RNP adaptors, which can act in their inactive form as an anchor [145], [146]; and (3) transmembrane anchors [26], [120].

The cellular environment of a mature neuron is a totally different situation. Due to the large distances mRNAs must travel and the complex spatial environments, active transport is required to keep the local transcriptome intact. In mature and active neurons, half of the mRNA population is motile. Most RNPs agitate in oscillatory movement with speeds up to 2μm/s. There is current evidence that RBPs (such as Fmrp) alter the speed and efficacy of RNP travelling in neurons. Deletion of Fmrp led to decreased travelling speed but did not influence the length of travelling, which indicates that Fmrp might be an adaptor for motor proteins [147], [148]. Why RNPs oscillate in neurons is not completely understood, but one explanation might be that this mechanism always ensures a fresh assortment of mRNAs at the point of requirement, preventing malfunctions caused by (1) old and damaged mRNAs or (2) mRNAs carrying outdated information in their modifications [93], [149]. The regulation of mRNA transport in neurons is thought to play a crucial role in neuronal plasticity [150]–[152]. Local translation is necessary for synapse-specific modifications and mRNA have to be stored locally as, for example, *Arc* mRNA, which is retained beneath spines in a UTR dependent manner [153]. The directed movement of different RNAs to different compartments is most probably mediated by a combination of localization techniques. RNAs usually do not move alone, as transport in batches is much more efficient. Furthermore, the predominant state of mRNA during transport is translationally repressed and they are activated or released upon stimuli [95]. The formation of these "batches", transport granule formation, is most probably a symphony of interplays between various RBPs, many mRNAs and adaptor proteins. They are held together via weak hydrophobic interactions between components of the granule,forming viscous droplets of some kind [126], [154], [155]. The weak interactions of those components enable a fast release and therefore de-repression upon the right stimuli.

Apart from cargo granules, mRNAs and proteins can also be transported in vesicles. On the one hand, the constant movement of vesicles through the cell along the microtubule cytoskeleton can be used to transport RNA anchored to transmembrane proteins, but on the other hand, vesicles can also inherit RNAs at their origin and deliver those at the place of cytosis [156]–[159].

Mechanisms of spatial degradation can also mediate RNA localization. This requires spatial recognition of the RNA via RBPs that mediate degradation or stabilization of the transcript. A prominent example, once more from *Drosophila melanogaster*, is the RBP staufen. During embryogenesis it is localized to the anterior pole of the oocyte. The staufen mediated mRNA decay pathway influences the local translatome. Stau1 binds Stau1-binding sites (SBS) within the 3'UTR of target mRNAs. During staufen mediated decay (SMD) the dsRNA binding is happening through either intramolecular dimerization of 3'UTR elements or by base pairing with lncRNAs (Alu elements) [137].

In neurons a prominent example is the Semaphorin3A (Sema3a) response system. Spatial translation in the neurite growth cone is crucial for the correct response to external stimuli. Local ubiquitination and degradation of RhoA, which is a mediator of Sema3a signaling, is continuously happening. Inhibition of RhoA degradation removes protein synthesis required for Sema3a-induces growth cone collapse. The local ubiquitination (Ngf induced and Smurf1 mediated) of newly synthesized proteins is major feature of local translation and growth cone response. Sema3a response needs continuous local translation and therefore induction of intra-axonal RhoA synthesis. Sema3a mediated growth cone collapse is reversible and can happen again after a short recovery time [160], [161].

1.3.3. Spatial translational control: Mechanisms of posttranscriptional gene regulation

To ensure proper intra- and extracellular functions, the molecular machinery must be regulated in a sophisticated way. Transcription is regulated by various transcription factors, DNA and chromatin modifications. Post-transcriptionally, RNAs are regulated by splicing, modifications and localization and their translation into proteins is regulated by numerous mechanisms. Most RNAs also play crucial roles in spatial protein production [162].

1.3.3.1. RNA editing

RNAs can be modified in multiple ways. Single bases or sequences can be exchanged, deleted or inserted. The editing machinery hereby involves guide RNAs for the correct recognition of target mRNAs. Most common is the substitutional editing. Adenosine or cytosine bases can be deaminated to Inosine and Uridine, respectively. Inosine is read as a Guanidine during the process of translation. These mechanisms contribute to protein diversity either directly via changing the coding sequence or indirectly by creating new or removing old splice junctions. Adar2 for example (Adenosine deaminase acting on RNA) edits its own pre-mRNA by changing an AA sequence to AI, which is read by the spliceosome as a new splice site (AG). This alternatively spliced isoform has a new translation start site and forms an altered protein

product. It was shown that exogenously induced RNA-editing can be used to include or exclude localization signals in endogenous RNAs. Recent studies focusing on endogenous mRNAs suggested that hyper-editing of the 3'UTR could influence transcript export from the nucleus and it has been suggested that A-to-I hyper editing drives association with p54nrb/NonO, which drives paraspeckle formation [163]–[165]. However, other studies showed that A-to-I hyper edited mRNAs localized preferably to the cytoplasm and that their nuclear export is not restricted [166]. In general, it has been shown that miRNA as well as RBP binding sites can be altered by Adar mediated A-to-I editing and therefore the fate of the mRNA is changed although the precise functional significance of 3'UTR editing remains unclear [167].

1.3.3.2. Nonsense-mediated mRNA decay (NMD)

Nonsense mediated decay (NMD) is a surveillance pathway that exists in all eukaryotes. Its main function is to reduce errors in gene expression. It eliminates mRNAs that contain premature stop codons, as the products of these mRNAs might be dangerous, since they could potentially have dominant negative functions over their canonical products, competitive functions, or completely new functions. The process of detecting aberrant transcripts occurs during translation. The first ribosome translating a message removes the exon-junction-complex (EJC) from the mature mRNA. If the stop codon is within 50nt from the last EJC or downstream from it, all components will be removed during the first round of translation and the transcript will be treated normally. When the ribosome leaves the mRNA prematurely due to a noncanonical stop codon, all downstream EJC will stay attach to the mRNA. Remaining EJC proteins are recognized by proteins of the NMD machinery. Termination of translation leads to the assembly of the termination complex composed of Upf1, Smg1 and release factors eRF1 and eRF3 on the mRNA. Upf1 can interact with remaining EJC proteins and especially Upf2 and Upf3, triggering phosphorylation of Upf1. Phosphorylated Upf1 interacts with Smg5, Smg6 and Smg7 [168]. Smg7 is thought to be the terminal effector in NMD as it accumulates in P-bodies, which are cytoplasmic sites for mRNA decay. In both yeast and human cells, the major pathway for mRNA decay is initiated by the removal of the 5' cap followed by degradation by Xrn1, an exoribonuclease enzyme [169]. It has been shown that NMD associated proteins are localized to axon growth cones and that NMD is important for axon guidance [170]. In spinal commissural neurons, which are a well-studied model for axon guidance, the axonal growth cone is initially attracted towards the midline floor plate [171]. However, the axon must undergo a drastic switch of responsiveness upon crossing the floor plate, since this is just an intermediate target and after crossing the floor plate the axon growth cone is repelled by the same stimulus. This switch in responsiveness is represented on the

molecular level as a change in expression of the axon guidance associated with receptor Robo3 [172]. It was shown that the two isoforms of *Robo3* (*Robo3.1* and *Robo3.2*) have opposing functions [173]. The isoform switch, however,is not obtained through induced alternative splicing but rather through fast degradation of the *Robo3.2* isoform due to a retained intron with a premature stop codon [174].

Furthermore, the *Arc* mRNA, which is located at the synapse, undergoes targeted NMD upon stimulation of a receptor (Nmdar). The mRNA contains a retained intron and, therefore, a premature stop codon but is not targeted by NMD due to its translational repressed state. Upon activation of the Nmda receptor the translation boosts and the mRNA becomes a target for NMD [170].

These examples show that NMD can be targeted and stimulus-dependent and can act as a mode of mRNA localization. Another mRNA decay pathway is the deadenylation from the 3' end, following message degradation. This process can be triggered by various mechanisms including noncoding RNAs like miRNAs.

1.3.3.3. RNA regulation by noncoding RNAs: miRNAs, siRNAs, piRNAs, snoRNAs and lncRNAs

RNA interference (**RNAi**) is a gene silencing process that can lead to either degradation of an mRNA or inhibition of its translation. Main players in this mechanism are small ribonucleotides named micro RNAs (miRNAs) or small interfering RNAs (siRNAs) that are partially or completely complementary to their targets. Due to their complementary nature they can bind their targets via base pairing and guide processive enzymes to the mRNA. **siRNAs** originate from double-stranded RNAs that are recognized and processed into smaller fragments of 21 to 25 nucleotides by Dicer, an RNAse III family member [175]. These small, double-stranded RNA pieces exhibit 2nt overhangs on both sides and are transferred to the RNA-induced silencing complex (RISC) simultaneously to their generation through Dicer-RISC association. The guide RNA, which is the strand with the lower free energy of binding at the 5' end is selected, bound by Ago and exposed while the passenger RNA is degraded. The exposed guide RNA then enables the sequence-specific targeting of mRNAs via complementary base pairing. Due to perfect complementarity, Ago proteins loaded with siRNAs act as endonucleases and cut the target mRNA [176].

piRNAs are another form of small regulatory RNAs. In contrast to siRNAs or miRNAs, their biogenesis is not regulated by RNAseIII type enzymes that convert the double-stranded

precursor into single-stranded, functional, small RNAs but they originate from single-stranded RNAs. piRNAs exist in the germ line (ovaries and testicles) and are a bit larger than miRNAs, typically 26-31nt. They are less conserved and do not expose any specific secondary structure. They were predicted to act on transposons, since most of them are antisense sequences of transposons. Some are essential for spermatogenesis in mice, such asthe *Miwi*, *MiwiII* and *Mili* associated piRNAs. They act primarily on methyltransferases. Some of them are also loaded onto Ago3 [177]–[179].

lncRNAs, long noncoding RNAs, are defined by their length (>200nt) and build another group of not translated RNAs. Recent studies, such as the FANTOM5 project [180], [181] have identified about 30.000 lncRNAs. On average they are lower in abundance than mRNAs in mammalian cells. Interestingly, they are highly tissue-specific (80% show tissue-specific patterns, whereas just 20% of mRNAs show tissue-specific patterns) and are conserved in sequence but not in transcription between species [182], [183]. There are numerous known methods of activity so far. Some examples include the regulation of gene-specific and nonspecific transcription, splicing, translation, siRNA directed gene regulation, X-chromosome imprinting (*Xist*) [184] and epigenetics (*Hotair*) [185].

Endogenous expressed, ~22nt large interfering RNAs are called **miRNAs**. They originate from own gene clusters, transcribed by RNA PolII, or from introns of mRNAs that form imperfect mRNA stem loop structures. These structures are recognized and excised by Drosha and Dgrc8. The resulting pri-miRNA, which is a ~70nt small double-stranded RNA stem loop, is exported from the nucleus and cytosolically cleaved by Dicer, which liberates the ~22nt double-stranded miRNA complex. The duplex of guide and transfer RNA is simultaneously delivered to the RISC complex and the strand with the thermodynamically less stable base pairing at the 5' end is selected as the mature miRNA and loaded onto AGO while the other strand is degraded. Like siRNAs, miRNAs can guide Ago or the complete RISC via base pairing to the target RNA, but in contrast to miRNAs they do not exhibit perfect base pairing over their complete length in animals. Important for target recognition is the seed region, nucleotides 2-7/8 of the miRNA, which often show perfect complementarity [186]. Ago loaded with a miRNA rarely shows endonuclease activity in animals due to imperfect binding. Targets are rather translational repressed by other components of the RISC and/or decapped and deadenylated, which leads to 5' to 3' degradation. It has been shown that miRNAs play a crucial role in sculpting gene expression in eukaryotes and exhibit temporally and tissue-specific expression profiles [187]. However, it remains unclear why their function is of crucial importance, since the effect of a

single miRNA is conventionally rather small (twofold). Most studies depict their necessity in the fact that most genes are targeted by multiple and sometimes different miRNAs. Several models of potential miRNA function have been proposed from classical switches to gene expression buffers and expression decoys, yet all have in common the fact that miRNAs guarantee precision and robustness to gene expression patterns [176]. Until today about 2500 miRNAs have been annotated and more than 60% of the human transcriptome have been predicted to be regulated via miRNAs, though it is still not clear to which degree translational repression or target destabilization contribute to this effect. There are examples known where translational repression is completely uncoupled from transcript degradation, but these are rather rare [188]. It is not completely understood how translational repression of targets is happening on the molecular level, although experimental evidence of recent studies points towards a repression of translational initiation rather than a repression of elongation. Recent studies have shown that some miRNAs are found spatially enriched in the axon and are therefore part of the spatial translational regulation machinery, causing differential localization of proteins [96], [189]. It was suggested that the sequence specific regulation of the spatial translatome is also controlled by miRNAs [96], [190]. Several nervous system functions rely on miRNA-target interactions such as dendritic branching [191], spine morphology [192] and axonal pathfinding [193]. The conditional knockout of Dicer in mice neurons resulted in reduced branching, increased spine length and axon tract abnormalities [194]. Mir-134 was found to have particular importance since it localizes to dendrites and regulates synthesis of Lim kinase 1 (Limk1), which controls dendritic spine development in hippocampal neurons [195]. Furthermore, it is important for the translational dependent growth cone attraction of *Xenopus* spinal neurons [196], [197]. It has been shown that RISC components are present in axons in vitro and in vivo [198] and recently a dataset of miRNAs in sympathetic neurons was published [199] as well as a dataset for axon specific miRNAs in the developing cortical system [200]. All together this data suggests a comprehensive function of miRNAs in the regulation of spatial translation.

1.3.3.4. mRNA m6A methylation: Writers, readers and erasers

There are about 150 different known RNA modifications. N_6-methyladenosine (m^6A) is the most abundant internal modification in mammalian mRNA. The methylation of rRNAs and tRNAs has long been known, but recent advances in antibody-based sequencing techniques have also revealed abundant m6A methylation in mRNAs [201]. m6A methylation is thought to be involved in tumor genesis, embryogenesis, circadian rhythm, viral replication, non-

coding RNAs, alternative splicing, translational regulation, cellular differentiation, disease and meiosis, sex determination and dosage compensation [202]. Components of the m6A methylation machinery colocalize with RNA PolII, which results in methyltransferation very early in the lifetime of an mRNA. It might represent an ancient mechanism of positively fine-tuning gene expression [203]. The m6A methyltransferase complex comprises two methylases, namely Mettl3 and Mettl14, as well as adapter proteins, including Wtap, Rbm15, and Kiaa1429. m6A methylation is reversibly removed by two demethylases: Alkbh5 and Fto. m6A methylation induces structural rearrangements and can inhibit Watson-Crick pairing. There are several examples known for functionally important m6A methylation of mRNAs. Most of those functional methylations are within the 5' UTR of transcripts. It has been shown that eukaryotic initiation factor 3 (eIF3) directly binds to m6A methyl groups and can recruit the 43S initiation complex and thereby bypass cap independent translation [204]. Also, Ythdf1 recognizes m6A methylations and promotes translation via recruiting the initiation machinery leading to cap independent translation. HuR, a well-established RNA stabilizing protein that binds to U-rich segments of 3'UTR and precludes miRNA binding, also recognizes m6A modifications. There is evidence that the spatial constriction between m6A and HuR sites govern the interaction of modified transcripts with HuR [205].

It has also been proven that m6A methylations can play a role in metastasis and cancer. *Malat1*, a metastasis related lncRNA in lung carcinoma, utilizes the m6A switch to recruit Hnrnpc [204]. Also, methyltransferase like Mettl3 and Mettl14 and RNA-binding motif protein 15 (Rbm15) are all highly expressed in myeloid leukemia compared with other cancers. Their alterations are linked with poor prognosis, suggesting that elevated m6A levels predispose to this type of cancer.

Additionally, m6A enhances functions of the lncRNA *Xist* for X-chromosome inactivation [203]. Likewise, m6A has been proven to play a role in the regulation of alternative splicing, although it appears not to participate in localizing splice sites in large introns. m6A methylations within alternative exons induce their inclusion via direct recruitment of the m6A reader Ythdc1 and downstream splicing factor Srsf3. This restricts binding of the exon-skipping factor Srsf10 and promotes inclusion of the exon [206]. The loss of the m6A eraser Alkbh5 increases overall m6A methylation and results in more cytoplasmic mRNA, suggesting more rapid nuclear export of m6A marked mRNAs or higher stability of them [207]. The knockout of the RNA methyltransferase Mettl3 additionally altered poly-adenylation site selection, indicating a role of m6A methylation in alternative poly-adenylation [208]. During

stress response, for example due to ultraviolet radiation or heat shock, 5' m6A methylations are induced and cap independent translation of specific transcripts is upregulated. This mechanism involves the nuclear re-localization of Ythdf2 to prevent m6A removal by the Fto demethylase [209].

m6A has also been implicated in regulating mRNA decay. When transcription was paused with actinomycin D in cells lacking the m6A writers Mettl3 and Mettl14, the decay of many mRNAs was decelerated. The Ythdf2 protein recognizes m6A methylation via its YTH domain in mammals and directs target mRNAs to cytoplasmic P-bodies by recruiting the CCR4-NOT deadenylation complex, initiating deadenylation and degradation of methylated transcripts [209].

m6A effects on miRNA processing have also been published, where m6A methylation seems to be required for processing pri-miRNAs via binding of Dgcr8 and recruitment of Drosha [210]. It was suggested that m6A methylation might play a role in localization of target mRNAs, though this could not be experimentally detected so far.

1.3.3.5. Translational initiation and cap-binding proteins

The initiation of translation is a complex and highly regulated mechanism in eukaryotes. It requires at least 12 initiation factors (eIFn). It begins with eIF2:GTP, which escorts the initiator tRNA Met-tRNA^{iMet} to the 40S subunit. Together with several other initiation factors they form the 43S preinitiation complex. Initiation almost always occurs at the first AUG after the m7G-cap. This AUG is embedded in the consensus sequence "GCCRCCAUGG". eIF4F, a heterodimeric complex of eIF4E, G and A, binds to the m7G cap through the cap binding protein. Now eIF4G also binds PABP and circularizes the mRNA. This is thought to prevent initiation on broken transcripts or to allow faster reinitiation. The eIF4F-mRNA complex is subsequently joined by other initiation factors and the whole RNP is bound by the 43S preinitiation complex. Subsequently the ribosomal subunit scans the mRNA until it encounters the first AUG and forms the 48S initiation complex. The recognition is achieved mainly by base pairing with the bound Met-tRNAiMet. eIF2 catalyzes hydrolysis of its bound GTP and thereby induces the release of all initiation factors. The 60S subunit joins in a GTP dependent reaction, catalyzed by eIF5B, and forms the 80S translating ribosome, which is now able to decode the RNA code into a chain of amino acids [7].

In addition to the canonical cap-mediated initiation of translation, ribosomes can bind and translate mRNAs internally thanks to so-called internal ribosomal entry sites (IRES). These are complex secondary structures of mRNAs that can recruit ribosomes internally without the need of the cap structure. They were first identified in virus mRNAs, but recent studies showed that

also endogenous eukaryotic mRNAs can inherit IRES structures that are active under certain conditions. They provide another layer of translational regulation as they are not affected by changes in the canonical initiation machinery and can therefore provide robust translation under adverse conditions [211], [212].

1.3.3.6. mRNA structure and stability

The structure of a pre-mRNA or mature mRNA does have an impact on all aspects of its life. Secondary structures can be recognized by RBPs, can inhibit or promote translation, can inhibit or promote splicing and can also influence the stability of the transcript. This means that information is not only stored in the sequence of the transcript but also in the structure. This allows more information to be stored in the genome without increasing its size [213]. Structural analysis of mRNAs has been and still is very complex and often rather approximative, since the structure depends the symphony of all forces acting on the polynucleotide. Since the molecular environment of the cell is extremely complex and dynamic and therefore not easy to model, most studies are either in silico or based on *in vitro* experiments under controlled conditions. Numerous forms of secondary and tertiary structures are known though the genome wide knowledge about structures was limited due to low throughput RNA structure probing techniques. Recent advances in structural analysis were made by methods such as fragmentation sequencing (FragSeq [214]), parallel analysis of RNA structure (PARS [215]) and selective 2-hydroxyl acylation analyzed by primer extension (SHAPE [216]), giving first insights into the endogenous RNA structure landscape on a genome-wide level. These techniques highlight the diversity and complexity of transcriptome and show on the other side that this research field is still in the early stages. It is now appreciated that structure can have influence on all classes of RNAs and throughout the whole lifetime [217].

The known influence of RNA structure on alternative splicing is mainly to bring together splice sites, expose them or mask them. For example, the human growth hormone GH1 is expressed in two isoforms with a big difference in expression. This effect relies on the fact, that the splice site of the lower abundant isoform is hidden by a favored stem loop structure. Also, the survival motor neuron protein Smn exists in two isoforms (*Smn1* and *Smn2*) which differ in inclusion of one exon (7). This gene does not change the coding sequence, rather the splicing pattern, by introducing an inhibitory element to the 3' splice site in *Smn2,* which results in skipping of exon 7. The 5' splice site is weak and hidden in a stem loop structure and strengthening the splice site leads to the exclusion of E7, while weakening it leads to its inclusion [218]. It has

been generally shown, that alternative splice sites have a higher GC-content than constitutive or skipped ones, which might suggests more stable secondary structures [217].

Ribosomal influence on mRNA structure is evident. The translating ribosome on the mRNA can penetrate secondary structures up to a certain mechanical strength [219], although these structures do alter the speed of translation [220]. Hairpins upstream of the initiation site can also inhibit initiation. This effect is most pronounced when the hairpin is near the G-cap. It was shown that the GC-content in the 5' UTR changes translation rates [221]. Secondary structures downstream of the initiation site can facilitate recognition of weak initiation codons. Secondary structures in the 3'UTR also alter translation. Structures can be recognized by diverse proteins. One prominent example is the 3'UTR of *Tf* which forms an iron response element and is bound by iron regulatory proteins when intracellular iron is low [222].

1.3.3.7. uORFs, alternative ORFs and overlapping ORFs

Approximately 40% of all mRNAs in mice exhibit alternative open reading frames [223]. These can either be upstream from the canonical reading frame or even overlapping with the canonical open reading frame. In general, they repress translation of the main ORF [224]. The probability of having a uORF increases with the increasing length of the 5' UTR, although uAUGs and uORFs are less frequent than statistically expected, underlining their negative selection pressure [225]. The ribosome sometimes skips the first AUG if it's too near to the 5' end or if the sequence environment is not recognized. If they are recognized by the 40S ribosomal subunit, however, they will downregulate translational efficiency of the following ORF. Once again, the secondary structure is of high importance, as is the distance to the Kozak sequence [226]. Known uORFs differ in length, number, distance from the cap structure, whether it overlaps the ORF or not and distance between the uORF-STOP and the main ORF START. Eukaryotic ribosomes can reinitiate after termination. This is the rate limiting effect of uORF-altered main ORF translation. The time for translation of the uORF is also important for the reinitiation rate. Reinitiation rates are difficult to study, since discrimination of reinitiation and leaky scanning is rather complicated. The best-studied example in human is maybe Atf4, which has two uORFs. uORF1 is just three amino acids (aa) in length, whereas uORF2 is 59aa long and overlaps with the main ORF. Under normal conditions (eIF-TC abundant) uORF1 is efficiently transcribed, and scanning 40S subunits after uORF1 will acquire eIF2-TC before reaching uORF2 and will reinitiate. Since uORF2 overlaps with the coding frame and is also longer than uORF1, it inhibits translation of the main ORF. During stress conditions, eIF2-α subunit phosphorylation is increased, which reduces the eIF2-TC

concentration. The 40S subunit that resumes scanning due to lack of eIF2-TC will skip the uORF2 and eventually reinitiate at the main ORF [227].

Inhibition or stalling uORFs often also triggers NMD, since the stalled ribosomes on the ORF cannot knock off the splice junctions proteins on the canonical ORF [228].

1.4. Methodical evolution in studying RNA localization in neurons: From the FISH to the mouse TRAP and beyond

Studying RNA localization evolved tremendously during recent years. Starting in 1969 [229] with intricate, *in vitro*, low throughput and extensive cost technologies like *in situ* hybridization (ISH), fluorescence *in situ* hybridization (FISH) and immunohistochemistry (IHC) based experiments modern, next generation sequencing (NGS) based methods allow the user to study the localization of thousands of transcripts at the same time. In general, methods can still be divided into low and high throughput-based technologies, each of which has advantages and disadvantages.

Low throughput-based techniques like the aforementioned ISH and IHC are restricted to few transcripts due to the nature of microscopy, yet enable the user to gain detailed information about specific transcripts. Advances in single molecule FISH (smFISH [230]) provided insights into the journey of single RNAs through the cell, while the usage of sophisticated reporter constructs allows us to study the translational state of the transcript in parallel [231]. All of these techniques suffer the same limitations however, which include (1) resolution limitations of the optical system, (2) detection limits of the camera system and resulting size of the reporter and (3) robustness of the biological system under imaging conditions. Initial experiments were conducted using radioactive reporter probes in ISH [229]. This has become a valuable method for studying the localization and expression levels of single nucleic acid sequences in a tissue, cell or even subcellular compartment. It is used to identify changes in expression and localization during embryonal development or cancer genesis. It detects a specific sequence through a hybridization procedure with a labelled complementary sequence in a fixed tissue or cell sample. Radioactive labels were first replaced by chemical reactions and have now been exchanged for fluorescent labels, giving the assay the catchy name fluorescence in situ hybridization (FISH [232]). Numerous technical advancements in probe engineering and the protocol improved the resolution, specificity and speed of FISH, making it irreplaceable in modern cell biology. In 1998 Singer and colleagues improved the resolution to the single-molecule degree, making the detection of single RNA [230] molecules in a fixed tissue or cell possible. The usage of multiple short probes per transcript, each coupled to a fluorophore,

amplifies the signal to the level of detection of modern light detection techniques. Single molecule FISH is now a powerful tool to study the localization and transcriptional state of specific transcripts in single cells, although it is still limited to relatively long and few transcripts per experiment and fixed samples, therefore providing just a snapshot of RNA localization.

Microinjections of labeled mRNAs into live cells enabled the tracking of mRNA trafficking live in neurons for the first time. Reporter RNAs are transcribed *in vitro* in the presence of fluorescently labelled UTPs. These mRNAs are introduced into living cells using an injection needle with a pressure-injection system. Although it is tedious and time consuming, this method gave first insights into the intracellular movement of mRNP particles [233], [234].

The introduction of **endogenously expressed mRNA reporters** in combination with target-specific, fluorescent and endogenously expressed fusion proteins enabled Bertrand et al. in 1998 for the first time to circumvent microinjections and simultaneously reach molecular resolution [235]. Multiple systems advanced from this point, all of which are based on fluorescent proteins fused to bacteriophage capsid proteins such as MS2, PP7 or λ in combination with multiple bacteriophage specific stem loop structures in the reporter [236], [237]. Thanks to positioning of the stem loops within the reporter and usage of multiple fluorophores it is now possible to track the journey of a single mRNA molecule live in a cell and simultaneously visualize the first round of translation of said reporter [231].

All of these techniques are still limited to single or few target mRNAs or reporters. Recent milestones in **transcriptome analysis** evolved RNA sequencing methods to a point at which it has become an inexpensive and fast way to study transcript abundance in a specific sample. Single cell-based techniques such as single cell sequencing and single cell real time PCR provide valuable information about variations in transcript abundance within cell populations and tissues, but are restricted to complete and dissociable cell types. Combinatory experiments of microdissections, cell organelle separations or cell compartment separations and subsequent transcriptome profiling is now the only way to study RNA localization on a global level. Limitations are due, of course, to the nature of the experiments: (1) the resolution and accuracy of the physical separation of subcellular compartments; (2) the physiology of the biological sample; and (3) the resolution on the cellular level. There are several published methods for the fractionation of subcellular compartments: (1) microdissection, followed by density centrifugations [238]; (2) microdissections, followed by affinity purification of cell specific RNAs [109]; and (3) cell culture-based techniques [152]. However, the cellular resolution of

most of these techniques remains rather poor, since the input demand of modern transcriptome profiling techniques in combination with low RNA yields of subcellular fractions usually requires the pooling of multiple cells. The homogeneity of the cellular sample is, therefore, a big challenge for all experiments.

1.4.1. In vivo systems to study intracellular localization in neurons

There are several published *in vivo* approaches focusing on subcellular RNA localization. Aside from giant cells like developing *Drosophila* egg cells, neurons are the most practical and widely used systems. A neuron can be divided into multiple fractions, but the most common physiological fractions are the dendrites, the axon, the cell body with the nucleus and the synapse, although the synapse is an intercellular structure, since it consists of a presynaptic nerve end, the synaptic cleft and the postsynaptic part. The accuracy of subcellular fractionation and purity of the test samples varies dramatically between the different approaches.

1.4.1.1. Microdissection followed by sucrose, Ficoll or Percoll density gradient centrifugation

The mouse brain is divided into at least 26 different physiological areas, all characterized by specific neuronal and nonneuronal cells [239]. The number of different neuronal and nonneuronal cell types varies from publication to publication and recent single cell sequencing techniques, in particular, have revealed that there might not be such a clear separation between cell types since most gaps are filled with innumerable subtypes and variations within one homogenous appearing cell populations might be tremendously [240]. Therefore, the isolation and biochemical analysis of subcellular fractions has to start with the purification of a homogenous cell population. Physiological dissections of brain areas under the microscope are called microdissections. Centrifugation allows brain microdissections to be further homogenized and synaptosomes to be isolated. Synaptosomes are isolated nerve terminals that encapsulate upon mechanical fractionation of the neuron. They consist of the presynapse, which contains the presynaptic vesicles, mitochondria and a postsynaptic density. Synaptosomes can be isolated via density purification. The first synaptosome purifications were generated in 1958 by Hebb and Whittaker [238] using the complete brain, but recent publications all use prior purification of cell fractions [239], [241], [242]. The homogenization is generally done via sheer force. Crude enrichments of synaptosomes can be achieved by pelleting myelin and other debris, transferring the supernatant to another tube and pelleting the microsomal P2 pellet at high speed. Even purer preparations are obtained by applying this crude synaptosome fraction to a density gradient centrifugation, using either sucrose [238],

Ficoll [243] or Percoll [244] gradients. Purified synaptosomes contain enough material for biochemical assays, maintain metabolic activity and membrane potential, and can still be stimulated to release neurotransmitters [245]. They are also used for further separation of synaptic vesicles, presynaptic cytomatrix and the postsynaptic density. Among impurities of the initial cell populations, the purification of synaptosomes tends to include impurities from mitochondria, the ER and other large cellular organelles. Recent publications revealed the sophisticated transcriptome [92], [246]–[249] and proteome of synaptosomes [250]–[252].

1.4.1.2. Microdissection coupled to NGS

Advances in microdissection techniques and detailed knowledge about physiological areas of the brain and the containing cell types have lead to various transcriptome analyses of simple brain slices. The challenge in these experiments is to dissect a homogenous sample consisting of just one cell type and subdividing this sample into fractions that can be analyzed via sequencing or microarray techniques. Reproduction of microdissecting tissues enriched for dendrites coupled to subsequent microarray analysis of the RNA content by different groups appeared to be impossible[239], [241], [242]. In 2012 Cajigas et al. published a deep RNA sequencing-based approach to analyze the transcriptome in the synaptic neuropil [246]. They micro-dissected the hippocampal neuropil and isolated fractions enriched for either CA1 neuron cell bodies or CA1 neuron dendrites. Comparing the extracted RNA pools between those fractions they identified 8,379 mRNAs associated with dendrites or axons in the hippocampal neuropil and confirmed their results via imaging of several candidates in primary hippocampal neurons. Due to the impurities of the microdissection they used a massive *in silico* filtering approach to identify the neuropil associated genes and filter out genes that might belong to glia, interneurons, blood vessels, nuclei or mitochondria [246]. Enriched genes in inhouse and published datasets of glia cells, astrocytes and oligodendrocytes [253], [254] were subtracted from the neuropil transcriptome. Also a combined list of interneuron specific transcripts from several publications [254]–[256] was filtered out. Finally mRNAs enriched in blood vessels [257] and mitochondria as well as transcripts coding for nuclear proteins were subtracted, leaving a list of 2,550 mRNAs that were annotated as the axonal and dendritic transcriptome. This pool of mRNAs includes transcripts for most biological functions and extends other previously published datasets by a factor of 25. This data suggests an enormous potential for spatial protein synthesis in the neuropil being a main player of the local proteome, although the massive filtering approach and the type of sequencing used causes an obvious loss of information. Subtracting enriched genes from interneurons, glia cells, astrocytes, oligodendrocytes and blood vessels may conceal many genes that are also part of the axonal or

dendritic transcriptome. Also, the subtraction of nuclear associated transcripts, though it might seem logical, bares the risk of losing many transcripts that either have an additional function besides the nuclear one or are translated in the distal part of the cell, though they function in the nucleus. Examples for the distal translation of proteins baring an NLS [258] or even the axonal translation of transcription factors like the cAMP response element binding protein Creb, coupled to retrograde transport [259], were previously published and are naturally, lost due to this massive filtering. The experimental setup is also insufficient to catch non-polyadenylated transcripts such as circular RNAs, miRNAs or other noncoding transcripts. Finally, the experimental setup, together with the massive filtering, enables a snapshot to be taken of the neuropil transcriptome, leaving out the question of which transcripts might be intracellular enriched in the spatial compartment compared to the somatic compartment.

Another type of microdissection coupled to deep sequencing was published in 2014 by Junker et al. [260]. Using massive cryotome histology coupled to low input RNA sequencing, followed by mathematical image reconstruction, they published a high-resolution 3D atlas of gene expression of a *Danio rerio* embryo. Cryo-dissecting microslides and extracting RNA for deep sequencing was performed here on a multicellular level but in theory this experiment could also be performed on huge cells like *Drosophila melanogaster* egg cells. The microdissection must be performed multiple times from different directions to enable mathematical image reconstruction, which limits this technique to uniform samples that can be aligned in silico after analysis.

1.4.1.3. Microdissection followed by affinity purification techniques coupled to NGS

In 2015 Ainsley et al. published a new method for collecting *in vivo* dendritic mRNA from mouse hippocampal CA1 pyramidal neurons [151]. They circumvented the challenge of physically or computationally separating dendrites from glia, astrocytes, blood vessels and interneurons by endogenous tagging of ribosomes in pyramidal neurons. This was accomplished by inserting an artificial fusion gene coding for the ribosomal protein L10a and GFP under Tet^{+}-control. This fusion gene was crossed into a mouse strain with Tet-on expression under the pyramidal neuron specific promotor Camk2a, leading to a cell type specific expression of the fusion gene. Tissue punches of the CA1 layer containing dendrites of the pyramidal neurons or the soma of CA1 neurons were analyzed. Crude lysate extraction via homogenization and immunoprecipitation of tagged ribosomes with associated mRNAs enabled them to explore the spatial translatome of pyramidal neurons. They identified 1,890

mRNAs to be bound preferentially by ribosomes in the dendrites of activated CA1 neurons. This finding highlighted the importance of local protein synthesis in dendrites to build the spatial proteome and, furthermore, the importance of mRNA localization in neurons. They collected samples at several timepoints of intensive synaptic activity (e.g. rest state and fear conditioned mice) and identified differences in ribosome associations upon synaptic activity. Dendritic RNAs appear to associate rapidly with ribosomes following a novel experience, providing strong support for their functional importance. Though it was shown *in vitro* before [261], this study showed *in vivo* translatome changes upon synaptic activity for the first time. Genes involved in translation and associated to cytoskeletal functions were most enriched in this condition. As the experimental setup allowed them to circumvent the massive filtering approach used by Cajigas et al., they were also able to identify enrichments for mRNAs encoding for proteins associated with nuclear functions such as chromosome organization (e.g. histone 4) and transcriptional regulation (e.g. transcripts of the mediator complex). In total 848 out of the 1,890 detected transcripts were also detected in previous studies but were filtered out due to association with (1) nuclear functions, (2) glia specific expression, (3) interneuron specific expression, or (4) blood vessel specific expression.

This finding agrees with models of synaptic plasticity that predict local stimulus sensitive translation as a main player of memory formation. This phenomenon was demonstrated in previous studies in synaptosomes [262] and *in vitro* cell culture [261]. Here, they provide the first *in vivo* experimental evidence for the synaptic triggering model, which is widely used to explain memory formation on the molecular level of the synapse. It proposes that an "activated synapse" adopts a protein synthesis independent "tagged" state that enables the neuron to capture plasticity related products (PRPs), such as mRNAs, that are translated in dendrites and proteins that are synthesized in the soma. The recruitment of these PRPs to the synapse stabilizes structural and functional changes within the synapse and enables long-term storage of memory. Synapses that are not activated will not capture PRPs required for the stable state. This model presumes a pool of local mRNAs in a translationally repressed state and a translational boost upon a specific stimulus. Cytoskeletal remodeling is, of course, a main requirement for structural changes in the synapse.

In 2016 Shigeoka et al. published a similar approach to catch snapshots of the axonal translatome in developing neurons [109]. They developed a similar method called Axon-TRAP-Ribo-Tag [263], which relies on the endogenous tagging of cell specific ribosome. In contrast to Ainsley et al., they used a Rpl22-HA fusion protein, which is specifically expressed in retinal ganglion cells (RGCs). The aim of this study was to capture changes in the axonal

translatome during embryonal development. Therefore, samples from retinal ganglion axons taken at three timepoints during the embryonal development (E17.5, P0.5, P7.5) and in the adult mouse were compared to the somatic fraction extracted from the retina. The total number of axonally translated mRNAs was higher in early stages of development, peaking at P0.5, and decreased postnatally. In contrast, the somatic translatome was more stable. Previous studies demonstrated that proteins are synthesized in developing axons, although it remained controversial as to whether mature axons exhibit local protein synthesis, also because previous studies found few or no ribosomes in mature axons [264]. However this was refuted by several studies that do indicate that ribosomes exist in mature CNS axons [265]–[267] and that their number is dynamically regulated [268].

This study proved for the first time the presence of active ribosomes in mature axons in the adult CNS. Interestingly, this study showed that only a subset (694 out of 2,576) mRNAs was spatially translated at all timepoints, revealing a high variability in translational regulation and perhaps as well in mRNA localization during neuron maturation. The study did not find transcripts associated with nuclear functions such as components of chromatin arrangement in axons of RGCs. Together with a similar finding in the *Drosophila melanogaster* visual system [269] this study shows that the maturation of presynaptic terminals is regulated by tightly controlled spatial protein synthesis.

About 15% of the ribosome-bound mRNAs might not undergo translation but are translationally inhibited. The retinal samples, which are used as somatic controls, unfortunately contain short axons and dendrites of the intra-retinal circuitry as well as their cell bodies. The expression of tagged ribosomes in non-retinal cells might also influence results. Finally, this study, like previous studies, focuses on just one small part of the spatial molecular environment; without detailed knowledge about the local transcriptome and proteome it is impossible to say whether ribosome association with local mRNA pools change or whether the pools themselves change (e.g. different mRNAs get localized).

1.4.2. Cell culture-based systems and their usage in investigating RNA localization

Cell culture-based systems produce more homogeneous cell samples. Results are not always easy to transfer to *in vivo* systems, but many general mechanisms are the same in cells growing in a complex tissue or in a culture dish. Also, they are way easier to manipulate. The origin of cells is of high relevance in this context. Cell lines that are cultured for decades tend to accumulate genetic defects and are not likely to reflect any endogenous cell type anymore. Additionally, most available cell lines are either genetically immortalized to regain

proliferating potential, are cancer cell lines, or must be cultured as stem cell lines and subsequently differentiated into the cell of interest. In contrast, primary cell cultures can be prepared by extracting cell material of the organism of interest and culturing these cells in conditioned medium. Of course, the extraction of just one cell type is not always easy, depending on the cell type of interest.

When studying neuronal localization, the cell system must be neuronal. The following chapter will discriminate between (1) immortalized cell lines, (2) neuroblastoma cell lines, (3) stem cell lines and (4) primary cell lines used to study mRNA localization *in vitro*.

Immortalized cell lines in this context are neuronal precursor cell lines that are extracted from the fetal organism and prevented from terminal differentiation via genetic modifications. A prominent example of these cells is the 'Lund Human Mesencephalic' cell line (LUHMES). This cell line was generated and characterized by the Leist lab in 2005 [270] as a model to study Parkinson's disease. It is a subclone of the human mesencephalic cell line MESC2.10, which originated from Lund university. The cell lines were immortalized by introducing a tetracycline-responsive *v-Myc* gene under control of the TET-off system. They can be easily cultured and proliferated. Upon tetracycline-induced downregulation of the *v-Myc* gene and simultaneous cAMP and GDNF treatment, they proceed in terminal differentiation and form a homogenous cell population of dopaminergic neurons, exhibiting biochemical and physiological dopaminergic markers and extensive neurite morphology. Functionally, they display spontaneous electrical activities and are able to release and take up dopamine, suggesting high similarity to human dopaminergic neurons [271].

Most common cell culture systems are based on endogenously "immortalized" cell lines, cancer cell lines or, in this case, neuroblastoma cell line [272]. There are numerous cell lines commercially available that all inherit the same background. Altogether they are defined by their origin from neuronal crest derived progenitor cells [273]. They do all have the potential to differentiate into neuronal cell types upon stimulus and keep proliferating potential *ex vivo* [274]. They are an easy and cost-effective test system for many biological questions and often used for drug screening systems for example for neurotoxins. Due to their origin and journey through many labs, most of them already carry multiple drug resistances and often also multiple mutations, peaking in polyploidy, making them not the best system for genetic cause and effect studies. Prominent examples of used neuroblastoma cell lines are PC12 (rat), SH-SY5Y (human), LAN-5 (human) or N1E-115 (mouse). Problems with these cell lines are the poorly

defined and controlled way of differentiation, the accumulation of genetic modifications and drug resistances, and their cancerogenic origin.

Stem cell lines such as induced pluripotent stem cells (IPS) or embryonic stem cell lines (ESC) are, as their names suggest, either dedifferentiated somatic cells that gained pluripotency, or extracted stem cells. Culturing these cell lines must be strictly controlled to prevent premature differentiation. Cell density, external stem cell transcription factors or inhibitory reagents must be used to keep those cells in their pluripotent or embryonal state. Due to their nature they have the potential to differentiate into various cell types, including neuronal cell lines, by adding external differentiation factors such as Brn2, Olig2, Zic1, Myt1l or Ascl1 [275], [276]. Of course, these factors can also be genetically integrated under a controllable promotor in the cell, which makes it a more defined and controlled system. The combination of differentiation factors, depriving stemness promoting factors, and media variations is crucial to generate homogenous neuronal cell populations.

The most sophisticated neuronal cell culture is the culture of primary cell lines originated directly from the organism of interest. These cells must be mechanically extracted via microdissection, enzymatically isolated, if possible and needed, sorted for specific cell types and then cultured under conditions enabling them to proceed in their endogenous differentiation program [277]. Impurities in the resulting cell line can therefore be the result of either imprecise microdissection or slight variations in the process of differentiation. Since the differentiation potential of stem cells decreases during their embryonal development, the timepoint of extraction is crucial for the generation of homogenous cell lines. Neuronal cells are, especially easy to damage during the process of microdissection and isolation due to their size and morphology. Common primary neuronal cultures include hippocampal, cortical and spinal cord cultures [277].

All these *ex vivo* test systems are used to study molecular localization. Some are more, and some are less like neuronal cells in the intact tissue, but all of them have advantages in usage for different biochemical assays. Since most neuronal cell cultures grow attached to a surface, they can all be used for microscopy and low throughput analysis of RNA localization. To gain material for high throughput biochemical analysis of subcellular compartments, these cells must be compartmentalized either mechanically or biochemically.

1.4.2.1. Microfluidic chambers

Microfluidic platforms can be used to isolate and direct the growth of CNS axons without the use of neurotrophins, providing a highly adaptable system to model many aspects of CNS neuro-degeneration and injury. In general, these systems consist of micro-vessels that are connected with micro-tubes of defined diameter. Liquid flow through these capillary vessels can be controlled by the experimental installation. Microfluidic platforms are perfect systems to study spatial manipulation of neurons. Due to their cost-intensive composition and regulation they are not suitable for high throughput biochemical approaches, although they are perfectly suitable for most forms of microscopy-based experiments. They enable extremely clean separation of subcellular compartments and, by varying the length and diameter of the tubes, also clear differentiation between axons and dendrites or neurons [152].

1.4.2.2. Campenot chamber

The Campenot chamber was the first device used in *in vitro* tissue culture to separate cell bodies from axons [278]. Campenot chambers are compartmented Teflon dividers attached to a collagen-coated petri dish via a thinly applied silicone grease layer. Typically, the nerve growth factor (Ngf) or similar growth factors promote neurite outgrowth through the grease layer and enable a clean separation between neurite outgrowths and somatic compartments that stays in the chamber [152]. The installation of these chambers is very fragile while the cell number is limited, making it not suitable for the generation of large amounts of biological material. Campenot chambers are traditionally used for spatial manipulations of neurons. They are more affordable than microfluidic devices and do not have to be regulated as precisely. Cortical, hippocampal and spinal cord neurons have so far not been successfully cultured in Campenot chambers, which further limits their application in localization experiments. It is mostly used to study target-derived neurotrophin-initiated signaling and retrograde signaling during axon development and synapse formation [279], [280].

1.4.2.3. Modified Boyden chamber

The Boyden chamber experiment was first introduced in the 1960s by Stephen Boyden [281] as an assay to examine mechanisms of cell migration. It consists of a two-compartment culture vessel separated by a porous membrane. Attractors on the outside of the membrane induce cell migration, which can be quantified by microscopy. Modified Boyden chambers are compartmentalized culture vessels, separated by a porous membrane of such a size that the cell cannot completely migrate through the pore. They are commercially available as cell culture inserts in numerous variations of size, material and pore size. Neuronal cells can be cultured on top of the membrane and outgrowing neurites can be attracted to the other side of the

membrane by a coating reagent. Traditionally, these coating reagents are parts of the basal lamina, such as laminin. Since the coating can be applicated by the user, these culture vessels are suitable for various cell types, including primary cell cultures. In contrast to microfluidic devices, the separated cell culture vessels can't be regulated as defined and these systems are not optimal for microscopy-based techniques due to the three-dimensional culturing. Nonetheless, modified Boyden chambers allow a clean, easy and cost-effective mechanical separation of subcellular compartments. In combination with a homogenous cell line, this culture system is perfect for studying molecular localization in neuronal outgrowths because by nature, it enables the user to extract higher amounts of material [81], [282].

2. Author contribution

Primary data analysis of RNA-seq, small RNA-seq experiments and K-mer enrichment analysis of the modified RESA experiment was performed by Vedran Franke (Akalin group, MDC-BIMSB bioinformatics platform). The RiboSeq data was generated by Camilla Cioli Mattioli (Chekulaeva group, MDC-BIMSB) and analyzed by Vedran Franke (Akalin group, MDC-BIMSB bioinformatics platform). The spatial LC-MS/MS data used for validation of the spatial SILAC data of Ascl1-iNs was generated by Alessandra Zappulo and analyzed by Koshi Imami. The primary analysis of the LC-MS/MS data from the N1E-115 cells was performed by Guido Mastrobuoni (Kempa group, MDC-BIMSB). The primary analysis of the LC-MS/MS data, as well as the SILAC and pSILAC data from the Ascl1-mESC and Ascl1-iNs, was performed by Koshi Imami (Selbach group, MDC-BIMSB). The primary analysis of the QuaNCAT data was performed by Erik McShane (Selbach group – MDC-BIMSB). The circRNA and lncRNA analysis on RNAseq data coming from Ascl1-iNs was performed by Andrei Filipchyk (N. Rajewsky group – MDC-BIMSB). The primary analysis of the modified RESA experiments and motif enrichment analysis was performed by Esteban Peguero Sanchez (Chekulaeva group – MDC-BIMSB).

Parts of this thesis were published in "RNA localization is a key determinant of neurite-enriched proteome" (Nature communications Zappulo et al. 2017).

3. Aim of the thesis

The aim of this thesis was to identify local protein and RNA pools in the used neuronal cell systems, to identify the impact of RNA localization on the establishment of local proteins pools and get insights into spatial protein production in neuronal cell systems. The second part of this work aimed on establishing a method for the *de novo* identification of zip codes in the used neuronal systems.

4. Material and Methods

4.1. Material

4.1.1. Equipment and consumables

4.1.1.1. Consumables for cell culture

Name	Manufacturer	Serial number
µ-Slide 8 Well ibiTreat	ibidi GmbH	9344
Cell counter slides	Luna	872010
TC 6-well plate	Sarstedt	50589
Sterile Filter 0.22µM 500ml	Life Technologies	TPP99505
Millicell Hanging Cell Culture Insert, P	Merck Chemicals GmbH	MCSP06H48
neoLab-Wattestäbchen Alu-Stiel 150 mm	NeoLab Laborbedarf	21022
Wattestäbchen Holzstiel 150 x 2,5 mm	NeoLab Laborbedarf	290121021

4.1.1.2. Consumables for biochemistry assays

Name	Manufacturer	Serial number
5&10% Bio-Rad Protean TBE precast gel	Bio-Rad	4565014 & 4565034
EPPENDORF DNA/RNA LoBIND S/L TUBES,1.5ML	Eppendorf	0030108051
Qubit Assay Tubes	Life Technologies	Q32856
Microseal 96-Well Skirted PCR Plates	Biorad	MSP-9601
Biosphere Fil. Tip 1000 blau	Sarstedt	70.762.211
Biosphere Fil. Tip 200 farblos	Sarstedt	70.760.211
Biosphere Fil. Tip 20 farblos	Sarstedt	70.1114.210

SafeSeal SurPhob Spitzen, 1250 µl, steril	Biozym Scientific GmbH	VT0270
SafeSeal SurPhob Spitzen, 100 µl, steril	Biozym Scientific GmbH	VT0230
SafeSeal SurPhob Spitzen,10 µl	Biozym Scientific GmbH	VT0200
Deckgläser für Mikroskopie 18 mm Ø, Nr. 1	NeoLab Migge GmbH	2637558
OBJEKTTRÄGER ISO8037/I GESCHN EINS.MATT	VWR International GmbH	631-0411
BRAND(TM) PCR-TUBES, 0.2 ML, COLOURLESS	Sigma-Aldrich Chemie GmbH	BR781305-1000EA
PCR-Gefäße, 8er, PP, ohne Deckel 0,2 ml	NeoLab Laborbedarf	104981325
PCR-Deckel, 8er, für PCR-Gefäße farblos	NeoLab Laborbedarf	104981334
RiboRuler High Range RNA ladder	Life Technologies	SM1823
40% Acrylamide/Bis Solution, 19:1	Bio-Rad Laboratories GmbH	161-0144
Quali - "Low binding" Microcentrifuge Tubes, 0,6ml	Kisker Biotech	G016
Quali - "Low binding" Microcentrifuge Tubes, 1,7ml	Kisker Biotech	G017
Quali - "Low binding" Microcentrifuge Tubes, 2,0ml	Kisker Biotech	G018

4.1.1.3. Equipment

Name	Manufacturer	Serial number
SP8	Laica	
SP5	Laica	
Trans Blot Turbo	Bio-Rad	1704155
Mini Trans-Blot Power Supply	Bio-Rad	1658033FC
CFX 96 touch qRT-PCR	Bio-Rad	10000068706
Mini Protean Tetra cell	Bio-Rad	1658003FC
Bioanalyzer 2100	Agilent	G2939BA
Qubit® 2.0 Fluorometer	Thermo Fisher Scientific	1104004846
Bioruptor	Diagenode	UCD-300
Gene Pulser II	Bio-Rad	165-2105
Luna automated cell counter	Logos Biosystems	CTM0011

4.1.1.4. Analysis – software

Name	Manufacturer	Version
CFX Manager	Bio-Rad	
LAS X	Laica	
R		3.3.3
RStudio		1.0.136
Serial Cloner	Serialbasics	2.6.1
Illustrator CS6	Adobe	16.0.0
nSolver	nanoString	2.6
Integrative Genomics Viewer		2.3.72

Fiji (ImageJ)	Wayne Rasband	1.48t

4.1.2. Chemicals and enzymes

Name	Manufacturer	Serial number
Western Blocking Reagent, 100 ml	Sigma-Aldrich	11921673001
FastAP Thermosensitive Alkaline Phosphatase	Fermentas Life Sciences	EF0651
PageRuler Plus Prestained Protein Ladder	Thermo Fisher Scientific GmbH	26619
Histone H3 Antibody-100 µL	Life Technologies GmbH	PA531954
Phusion High-Fidelity DNA Polymerase	Fermentas Life Sciences	F-530L
Alexa Fluor 647 Goat Anti-Guinea Pig IgG (H L)	thermo scientific	A 21450
Formaldehyde, 16%, methanol free 20x10ml	Polysciences Europe GmbH	2637557
Non-fat skimmed milk powder	AppliChem	A0830,1000
Urea, MB Grade	Merck Millipore	66122-500GM
CoverGrip Coverslip Sealant	Biotium	23005
TEMED Electrophoresis reagent, 25 ml GEB	Thermo Geyer	SA/T9281/000025
GeneRuler 1 kb Plus DNA Ladder-5 x 50 µg	Life Technologies GmbH	SMF-1082-5-BS
L-Glutamine, 200 ml Sterile filtered	Sigma-Aldrich Chemie GmbH	G7513-100ML
Laminin from Engelbreth-Holm-Swarm Murine Sarcoma Basement	Sigma-Aldrich Chemie GmbH	L2020-1MG

Membrane, Aqueous Solution, 1mg/ml		
Alexa fluor 488 Goat Anti-Chicken	Life Technologies GmbH	A11039
Alexa Fluor 488 Donkey Anti-Mouse IgG (H L) Antibody	Life Technologies GmbH	A21202
Poly-L-Ornithine Hydrobromide	VWR International	114-15
DMSO, Cell culture grade	Genaxxon Bioscience	M6323.0250
recombinant mouse LIF Amsbio	Amsbio	AMS-263-100
dNTP Mix, 10mM each	Thermo Fisher Scientific GmbH	11853933
BD Matrigel GFR-RED. Phenolfrei	Bioline	734-1101
prolong Gold Antifade Mounting with DAPI 10mL	Life Technologies GmbH	P36931
peqGOLD TriFast 100 ml DNA/RNA/Protein	VWR Internat. GmbH (Peqlab)	30-2110
RQ1 RNase-free DNase	Promega	M6101
Mini-PROTEAN TBE Gels, 5%, 30 microliters	Bio-Rad Laboratories GmbH	456-5013
Endura electrocompetent cells	Lucigen	60242-1
NuPAGE Novex 4-12% Bis-Tris Protein Gels, 1.0 mm, 10 well-10 gels	Life Technologies GmbH	NP0321BOX
Agencourt® RNAClean™ XP, 40 mL	Beckman Coulter GmbH	A63987
TGIRT™-III Enzyme	Ingex	TGIRT50

GSK-3 inhibitor	Merck Millipore	361559-5MG
StemMACS PD0325901 (10 mg)	Miltenyi Biotec GmbH	
GlycoBlue Coprecipitant (15 mg/mL)-5 x 3	Life Technologies GmbH	AM9516
Bio-Rad Protein Assay Dye Reagent Concentrate	Bio-Rad Laboratories GmbH	500-0006
Low Molecular Weight DNA Ladder	New England Biolabs	N3233 L
Penicillin/Streptomycin 10.000 E/10.000	Biochrome	A 2213
D(+)-Saccharose für die Molekularbiologie	AppliChem GmbH	A2211
Agencourt® AMPure® XP, 60 mL	Beckman Coulter GmbH	A63881
Tris(hydroxymethyl)aminomethane molecular biology grade	Serva Electrophoresis	37186.04
SDS für die Molekularbiologie	AppliChem	A2263,0100
RNase OUT	Life Technologies GmbH	10777019

4.1.3. Kits

Name	Manufacturer	Serial number
PLASMID MINIPREP DNA Purification Kit	roboklon GmbH	E3500-02
NUCLEOBOND XTRA MIDI PLUS, 1 PAK	Macherey Nagel	MN/00740412/000050
Lipofectamine messengerMAX	Thermo Fisher scientific GmbH	LMRNA008

Lipofectamine 2000 Transfection reagent	Thermo Fisher scientific GmbH	11668019
JetPrime	Peqlab	13-114-15
GeneJET Gel Extraction Kit 250 rxn	Life Technologies GmbH	K0692
Qubit dsDNA HS Assay kit 500 assays	Life Technologies GmbH	Q32854
Qubit RNA HS Assay kit	Life Technologies GmbH	Q32855
Qubit Protein Assay kit 500 assays	Life Technologies GmbH	Q33212
TruSeq® Small RNA Library Prep Kit	Illumina	RS-200-0012
TruSeq® Stranded Total RNA LT	Illumina	20020596
Bioanalyzer DNA1000	Agilent Technologies	5067-1504
Bioanalyzer High Sensitivity DNA Kit	Agilent Technologies	5067-4626
Bioanalyzer RNA nano	Agilent Technologies	5067-1511
Bioanalyzer RNA pico	Agilent Technologies	
Ribo-Zero Magnetic Gold Kit (Human/Mouse)	Biozym Scientific GmbH	191845
ERCC RNA spike ins	Thermo Fisher scientific GmbH	4456740
TranscriptAid T7 High Yield Transcription Kit-50 reactions	Life Technologies GmbH	K0441
mMESSAGE mMACHINE T7 in vitro transcription kit	Thermo Fisher Scientific GmbH	AM1345

Click-iT protein extraction kit	Thermo Fisher Scientific GmbH	C10416
Poly(A)-Tailing Kit	Thermo Fisher Scientific GmbH	AM1350
Phire Animal Tissue Direct PCR Kit	Life Technologies GmbH	F140WH
MAXIMA FIRST STRAND CDNA SYNTHESIS KIT	Thermo Fisher scientific GmbH	K1642
SensiFast SYBR no-rox 2000xreactions	BIOLINE	BIO-98020
Direct-zol RNA mini-prep kit	Zymo	R2052

4.1.4. Buffers, solutions and media

Media and solutions

Name	Manufacturer	Serial number
DPBS without Ca and Mg	PAN Biotech	P04-36500
10x Tris/Boric Acid/EDTA (TBE), 1 L	Bio-Rad Laboratories GmbH	161-0733
MEM Non-Essential Amino Acids Solution, 100X	Sigma Aldrich	M7145-100ML
100X Nucleosides, 50ml	Merck Millipore	ES-008-D
Neurobasal(R) Medium (1X), liquid	Life Technologies GmbH	21103049
Fetal bovine serum embryonic stem cell	Life Technologies GmbH	10828028
Dulbecco's Modified Eagle Medium (D-MEM)	Life Technologies GmbH	61965059

Opti-MEM I reduced serum Medium	Life Technologies GmbH	31985062
TrypLE Express (1X) Phenol red	Gibco	12605028
Knockout DMEM	Life Technologies GmbH	10829018
KnockOut Serum Replacement	Life Technologies GmbH	10828028
Fetal bovine serum	Life Technologies GmbH	10270-106
Advanced DMEM/F-12-10 x 500 mL	Life Technologies GmbH	12634028
SILAC Advanced DMEM/F-12 Flex Media	Life Technologies GmbH	A2494301
N2 supplement 100x, liquid	Life Technologies	17504044
B27 serum-free supplement 50x, liquid	Life Technologies	17502048
LB-Agar (Luria/Miller) 2,5 kg	Roth - Carl Roth	X969.3
LB-Medium (Luria/Miller) 2,5 kg	Roth - Carl Roth	X968.3

Cell culture media

N1E Growth medium (555ml): 500ml DMEM, 50ml FBS, 5ml Pen/Strep (100x)

N1E Differentiation medium: 500ml Neurobasal, 5ml Pen/Strep (100x)

PUCK's saline solution: 0,17 mM Na_2HPO_4, 0,22 mM KH_2PO_4, 20 mM Hepes, 138 mM NaCl, 5,4 mM Ca_2Cl, 5,5 mM Glucose, 58,4 mM Sucrose

2i medium (420.85ml): 200ml Advanced DMEM/F12, 200ml Neurobasal, 4ml L-Glutamine (100x), 4ml ßME (55mM), 4ml N2-supplement (100x), 8ml B27-supplement (50x), 50µl LIF (10^7 U/ml), 400µl GSK-3 Inhibitor (3mM), 400µl PD0325901 (1mM)

mESC medium (295.025ml): 250ml Knockout™ DMEM, 35ml ES-FBS, 2.5ml L-Glutamine (100X), 2.5 ml ßME (55mM), 2.5ml Non Essential Amino Acids (100x), 2.5ml Nucleosides (100x), 25µl LIF (10^7 U/ml)

Feeder's medium (277.5ml): 250ml Knockout™ DMEM, 25ml ES-FBS, 2.5ml L-Glutamine (100X)

AK medium (595.025ml): 250ml Advanced DMEM/F12, 250ml Neurobasal, 5ml L-Glutamine (100X), 2.5ml ßME (55mM), 75ml Knockout-Serum Replacement, 5ml Pen/Strep (100x)

Monolayer medium (104ml): 100mL Advanced DMEM/F12, 1ml N2-supplement (100x), 2ml B27-supplement (50x), 1ml Pen/Strep (100x)

mESC Freezing medium (10ml): 5ml ES-FBS, 3.2ml mESC medium, 0.8ml 2i medium, 1ml DMSO

4.1.5. Oligos

qRT-PCR oligos

gene	FW primer	REV primer
Bmper	GCCTGGGATTACCTGCTGC	ACACATTATGCAAGGGTTGTCTG
Camk2b	ACCACTACCTGATCTTCGACC	CCGCCTCACTGTAATACTCCC
Cdc42bpg	GGCCTTGGCTACTAAGATGGC	TTGGCACGGATCTCAGCTTC
Col3a1	CTGTAACATGGAAACTGGGGAAA	CCATAGCTGAACTGAAAACCACC
Crtap	CGCCGAACTACGCCTCTTC	TGCTTCAGGAGATAGGTGTGA
Dab2	CCTTCATTGCTCGTGATGTGA	CCCCAAACAAATCCATCTGGTC
Fbll1	CGAGGGCGTGTTCATCTACC	CTGCCAATTTGGAGCGGAAG
Gng3	GCACTATGAGTATTGGTCAAGCA	GTGGGCATCACAGTATGTCATC
H2AFY2	CGGTTGATGCGCTACTTGAAA	CAATGACCGCAGCCATGTAGA
H3f3a	TGTGGCCCTCCGTGAAATC	GGCATAATTGTTACACGTTTGGC
Igf2	CCGAGAGGGACGTGTCTAC	GTCTCCAGGTGTCATATTGGAAG
kif1c	GGAGCCTCCGTGAAAGTTG	CCGAAGTATGCGACCAGTAAGA
Lamb1	AGACCCGAAGAAAAGACAGGC	CCATAGGGCTAGGACACCAAA
ldlrap1	CTCAAGTACCTTGGTATGACGC	CTGTAAATGGACACGTTCTCGAT
Mbnl2	CCCAAAAGTTGCCAGGTTGAA	CTGGGTTTTTAAGTGTGTCGGA
Mme	CTCTCTGTGCTTGTCTTGCTC	GACGTTGCGTTTCAACCAGC
Mov10	GAGGTTCGAGAGTTTTCTGGC	TCGTGGTTGTAAATCGTCCGC
Myo1c-1	GAGCAACCCCGTGTTAGAGG	ACTTTTCCAGGAGGTAACTGAGA
Myo1c-2	ATGGAGAGCGCCTTGACTG	TCGGTAGGGATTGACAGAGAC
Nid2	CTCTTTCCTTACGGGGAGTCG	GGCATCGTAGAAACGCAGG
nucleolin	AAAGGCAAAAAGGCTACCACA	GGAATGACTTTGGCTGGTGTAA

Nup210	GTGACGCCATCGTTGATCTCA	GTCCAGTCGAAGACCAGTCC
Nxf7	ATGTGCTCTAATGAAAGGAAGCA	GCTCAGGGAAGAAACCCTGT
pura	ATCCGCCAGACAGTCAACC	TCCACTCCATAGTCGTCGATG
Rbm17	TGAAGGCTGGTCCAAAAACTT	CCGGAGCAAGCACTGTACT
RBM47	GCCATGAACAGCGATCCAAC	CCGGTGCGCTCTATCAGTG
18S rRNA	AAACGGCTACCACATCCAAG	CCTCCAATGGATCCTCGTTA
St3Gal6	GGGGAACAAATGGCTATTGGT	AGGGCAACGGAAAATTATTGGT
Stx3	GAAGGCACGGGATGAAACTAA	GGACAGTCCAATAATCAACGCTA
Tagln	CCAACAAGGGTCCATCCTACG	ATCTGGGCGGCCTACATCA
Tpt1	CATCAGCCATGACGAGCTGTT	CTCTGTTCTACTGACCATCTTGC
Tubb3	TAGACCCCAGCGGCAACTAT	GTTCCAGGTTCCAAGTCCACC
Vangl1	GATACCGAATCCACGTATTCTGG	TCTGCCATCTTTATTCCTTGGTG

RESA amplification oligos

RESA #	SEQUENCE_ID	PRIMER_LEFT_SEQUENCE	PRIMER_RIGHT_SEQUENCE
1	Kcnk12-ENSMUSG00000050138-ENSMUST00000055221	CTCCAGGTAGAGCGGCCG	AAGAAAGCGCCATCTCGGAG
2	Hspg2-ENSMUSG00000028763-ENSMUST00000030547	CTCATAGGCACCCACCTGC	TCCCTCTAGGACCCAAGAGC
3	Myo1d-ENSMUSG00000035441-ENSMUST00000041065	TGACTGCTGCACATCAGAGG	CAGTCACACCCCGAATTCCA
4	Hapln3-ENSMUSG00000030606-ENSMUST00000032827	AGGCCTAGACTGGTGCAATG	ATCCAAGCAATCCCCCAACC
5	Mfge8-ENSMUSG00000030605-ENSMUST00000032825	TAATGCTCAGTCCTGCCAGC	GTAGCAAGCCAGCAGAGGTC
6	Gm2399-ENSMUSG00000078141-ENSMUST00000104944	GGTCTATGTGAGTTGGGGCC	ACTAGCCAGCAAAGGGTGAG
7	Fam129b-ENSMUSG00000026796-ENSMUST00000028135	CCTCCAGGCTTCTGTTGGAC	ACCCCCAAGGCACTAGAGAA
8	Mov10-ENSMUSG00000002227-ENSMUST00000106774	TAGCAAGTTCAGCTGCCGAA	GCTTTCTTCCTCCCCCTTCC
9	Snai1-ENSMUSG00000042821-ENSMUST00000052631	CATCCTCGCTGGCATCTTCC	AAACATCTTTCTCCCGGGGC
10	Smtn-ENSMUSG00000020439-ENSMUST00000170588	AAGTCCTAACCCCTGCTTGG	CAAAACGCTGCGTGTGTACA
11	Tmem109-ENSMUSG00000034659-ENSMUST00000038128	AGAAGAGTGACAAGGGCGTC	ACTGCAGCCATGATTCCACA
12	Jph2-ENSMUSG00000017817-ENSMUST00000109425	CTACCAGTGCCAGGAGTTCC	TTGGCATCCTGAACCCATCC
13	Arhgap10-ENSMUSG00000037148-ENSMUST00000076316	CTCCTGTAGCTCCCGGCC	TTTACAGAGGGGCCTTGTGG
14	Stard10-ENSMUSG00000030688-ENSMUST00000164479	GGCTCTCTGCAAGGACCAAG	TGGGGCTGTGTGGTTCATTT
15	Gpx3-ENSMUSG00000018339-ENSMUST00000125094	GTACCACTGATGCCCCCAC	ACCCTCTCTCAGCCCTCAAT
16	Gaa-ENSMUSG00000025579-ENSMUST00000106259	GTCCTAGGAGAGTCCGTCGT	ACACTTCACAGTAAGTTTCTCGAGT
17	Pde1b-ENSMUSG00000022489-ENSMUST00000023132	GATTAGGGAACTGGGGGCTG	GGAGTCACATGCGGTTCAGA
18	Fkbp10-ENSMUSG00000001555-ENSMUST00000001595	TCTGAGGGCCAGAGTGTCAG	TTTGAGTCCAAGACCCTGCC
19	Dusp9-ENSMUSG00000031383-ENSMUST00000019701	CACATAAAGCCGCATGGAGC	CACAGGTATTGCCAGCTCCA
20	Tln1-ENSMUSG00000028465-ENSMUST00000030187	TTAATGCAGACCCAGCCCAG	TGGAGTCTGACCTGTTTGGG
21	Dnaaf1-ENSMUSG00000031831-ENSMUST00000093100	GACTGAACCCCAAGGCAGTC	GGATATTCCCGAGAGCTGGC
22	Emilin1-ENSMUSG00000029163-ENSMUST00000031055	ACAGGTGTAGACAGGGTGGA	TCCTGAATAGTTTAGTACGGTGGT
23	Arfgap1-ENSMUSG00000027575-ENSMUST00000029092	TAGGAGCCTAGCACCCTAGC	TGCACTGCTGGTCAAACTGA
24	Rab3il1-ENSMUSG00000024663-ENSMUST00000121418	CCCAGGCCTGAAGTAAGCTC	GCACGAGGCCAGTTAGGAAA

25	Itga5-ENSMUSG00000000555-ENSMUST00000023128	CCTGAGCCCTCCTGATCTCA	AGGGTTGGCTAGGGAGAGAG
26	Necap2-ENSMUSG00000028923-ENSMUST00000030760	GCACAACCTTTTCCTCAGCC	TCTGGTCACACAAGGTCAGC
27	Anxa6-ENSMUSG00000018340-ENSMUST00000108883	AGAGGACTAAAGCTCCTGACCT	CAAAGCTCAAGTCAGGCAGG
28	Csrp1-ENSMUSG00000026421-ENSMUST00000097561	CTCAGAGTGAGGCTGCCAC	TCAACAGCCCATGCCTACAG
29	Plod1-ENSMUSG00000019055-ENSMUST00000019199	CGATCCCTAACTGCCCAGTG	CAGTGACAGGTAACGCAGGT
30	B4galt5-ENSMUSG00000017929-ENSMUST00000109221	TGAGATGCGAAGAGAAGCCG	GCTCTGTGCTGTCCTCTGAG
31	Cdc42bpg-ENSMUSG00000024769-ENSMUST00000025681	CTCCCCCTGATGCCCTCT	TCTTTCACTCCCCACCCTGT
32	Itgb5-ENSMUSG00000022817-ENSMUST00000115028	TGGACTGAGGCTCCTGGATG	ACGATCAACTGGAAAAAGCCA
33	Tinagl1-ENSMUSG00000028776-ENSMUST00000030560	CTAGGCGAGGTGGGATCCA	AGTGAGGGCCTAGAGAAGGG
34	P2rx5-ENSMUSG00000005950-ENSMUST00000006104	GCTCTTGGCAGGGAACAGAT	CATCAAGGCTGGGCTTCTGA
35	Rcn3-ENSMUSG00000019539-ENSMUST00000019683	TGAGCTCTGATCCCCATGTC	AGGCAGGACAGGTTAGCTTG
36	Ldlrap1-ENSMUSG00000037295-ENSMUST00000037828	GGGGAGAGGGGGAGAAAGAT	CCCAAAAGCAGAGCTACCGA
37	Naga-ENSMUSG00000022453-ENSMUST00000023088	AGAATGACAGGCTACGGCAG	TAGCGACCTGAGGGAGCTAA
38	Lmf2-ENSMUSG00000022614-ENSMUST00000023283	AAATAGCGTCGTGTGCTACCA	ATGGAGAACATTGAGCTACAAGGA
39	Slc29a1-ENSMUSG00000023942-ENSMUST00000167692	CCTGTGGGGACAGAAGAACT	CCTCTGAAGGCACCTGGTTT
40	Wars-ENSMUSG00000021266-ENSMUST00000160477	CAGAGCAGCTGTCCAGTGAA	CCCTTGTGCCTTCAGAGAGG
41	Pdgfrb-ENSMUSG00000024620-ENSMUST00000115274	GACATCACTCCATTTTGCCCG	TCCGGTGTCTAAATGTGGGT
42	Kcng1-ENSMUSG00000074575-ENSMUST00000109191	GTCCTCCCTCTAAGCCCAGA	GGCTCTATCGCTTCTGGTCC
43	Lgi4-ENSMUSG00000036560-ENSMUST00000039775	AAAGACCCTGGAAGTGGCTG	GGGGTAACAGTCACGAGTGG
44	Ccdc134-ENSMUSG00000068114-ENSMUST00000089174	GAGTTGTAGCCCTGGACCAG	CAATGAAGGCACAGCACAGG
45	Myh9-ENSMUSG00000022443-ENSMUST00000016771	GAGCTTCTCCTGCAGCCC	ACTGTAGTGGTGACAGCTGC
46	Fxyd3-ENSMUSG00000057092-ENSMUST00000167369	CAGGTGCCAGAGACTCTGTG	TCAAGCCAACATGACCTGCT
47	Slc13a5-ENSMUSG00000020805-ENSMUST00000021161	TTAGGAAGAGCCGCAAGAGC	GCAAGCCTTCAACAGTGTGG
48	Egfl7-ENSMUSG00000026921-ENSMUST00000139801	AGAACAGAAACTGCAGTTGGTG	GAGTTCTTTCTCCGGGCTCC
49	Egflam-ENSMUSG00000042961-ENSMUST00000160207	GACATCAAGCTGGGAGCCTT	AGCTACAAAATCAGGGGGCC
50	Tead4-ENSMUSG00000030353-ENSMUST00000006311	TTGGAGAGCAATGGTAGGGG	ACATGTGGATGTGCTCAGCA
51	Myh6-ENSMUSG00000040752-ENSMUST00000081857	TAACCTCTCCAGCAGACCCT	TGGACTTGTGCCTAGCTGTG
52	Erp29-ENSMUSG00000029616-ENSMUST00000111802	GGCCCTTAGAGCTCTGGACT	GCCATCCTGTCCCTGTTTCA

53	Stat3-ENSMUSG00000004040-ENSMUST00000103114	ATCTTTGGGCAATCTGGGCA	ACTGGGCTGTGTGATGGATG
54	Slc12a4-ENSMUSG00000017765-ENSMUST00000034370	TATTCCTGAGCCCCGATGGA	GCTCCATTCCACTTGGCTTG
55	Mettl7b-ENSMUSG00000025347-ENSMUST00000026398	CTCTCCCCAAGGATGCCATC	GGCTGCTTTATTGAGTGCCG
56	Stra6-ENSMUSG00000032327-ENSMUST00000170397	AACCACATCACCCTGCCTTT	GGAATGTTCCTCAGAGGCACA
57	Plekhg3-ENSMUSG00000052609-ENSMUST00000075249	AAGGAGAGAACAGCCACGTG	GTGTGTCCATTGGCAACCAC
58	Des-ENSMUSG00000026208-ENSMUST00000027409	AGGAATTCAGTGTCCTGGCC	ACCCTTCCATTTCCCCTAACT
59	Tmem88-ENSMUSG00000045377-ENSMUST00000050140	TGGGTCTGAGGTGCTTAGAGT	GGTACCAGGACACCACATCG
60	Fgf3-ENSMUSG00000031074-ENSMUST00000155320	CACCTGGAAAGCCCTCTGAG	CAGTCCCCTATTCCTCCCCA
61	Hyou1-ENSMUSG00000032115-ENSMUST00000161318	TATAACCCCAGCTCCTGCCT	CCAACCAGCCTGAACTGCTA
62	Tagln2-ENSMUSG00000026547-ENSMUST00000111230	ACTCTCTCTCCTTCCCCTGC	AACTCGTGTTCGTCAAGCCC
63	Emilin3-ENSMUSG00000050700-ENSMUST00000057169	GGATCGATGCACAGAAAGCG	GGGCAAACACTGAATGCAGG
64	Pbxip1-ENSMUSG00000042613-ENSMUST00000038942	CCCTGCACTGGTACCACTTT	GACATGGAAACCAGAGGCCA
65	Hkdc1-ENSMUSG00000020080-ENSMUST00000020277	GGACATCTAGGTGGCCCAG	AATGCAGAAAACAGCGACACC
66	Yap1-ENSMUSG00000053110-ENSMUST00000173264	TCCGATCCCTTTCTTAACAGTGG	TATCTAGCTTGGTGGCAGCC
67	Ldlr-ENSMUSG00000032193-ENSMUST00000034713	GAGCCGTCTCTTTCCGGGAT	CTCAAAGCAACGCACAGGAC
68	Lrrc38-ENSMUSG00000028584-ENSMUST00000052458	AGATGACTGAGCCACCTCCT	GCTTCTATTCCCCCTCAGCC
69	Col4a2-ENSMUSG00000031503-ENSMUST00000033899	GAATCTGTGAGGCTCCCGC	TGACACAGGTTGTAGGACGC
70	P4hb-ENSMUSG00000025130-ENSMUST00000026122	TGTAGTGCAGAAGCCAGATCTG	TTCATTCAGCCCCAGTGTCC
71	Lrpap1-ENSMUSG00000029103-ENSMUST00000030986	CTGATCCTACAGTGGCTGGC	TGTGAACCCATGTGCCGTAG
72	Chst15-ENSMUSG00000030930-ENSMUST00000077472	GAGCTGAGTCCCTTCTGCAG	TCACCCAAAACACTGCCTGT
73	Cndp2-ENSMUSG00000024644-ENSMUST00000025546	CAGTTTCCTCGGAGTGCTGT	TAGCTGCAACTGGACTGGTG
74	Hdlbp-ENSMUSG00000034088-ENSMUST00000170883	AGAAACCTTTCCAGCCTGCT	GCCCTAGGCTCCTCAACATC
75	Susd2-ENSMUSG00000006342-ENSMUST00000095541	AGAGCACACTTGGCTGACAG	GGCAGGGCTCAAACTCTGAT
76	Syne3-ENSMUSG00000054150-ENSMUST00000095439	GCACAGGTGACCTTCTGCA	AGACACGCCCACATTCGTAG
77	Adamtsl4-ENSMUSG00000015850-ENSMUST00000117782	TTGAAACGGGTCCCAGAGAC	CTTCTGGGGTGGGCATCATT
78	Fbxo2-ENSMUSG00000041556-ENSMUST00000047951	CCCTGAACGACCTCCCCTA	ATTGAGTGCCCACTTTGAGG
79	Pdzd3-ENSMUSG00000032105-ENSMUST00000034618	AGGACATCCCTGTTCAGCCT	TAAAGTGCTCACAAAGAGGTCAGA
80	Flnb-ENSMUSG00000025278-ENSMUST00000052678	GCCAATGCTGGACGTGTCC	CGGTCTCCGTGTCTATTCCG

81	Adam19-ENSMUSG00000011256-ENSMUST00000011400	TTCTTACCCGAGTGCCTCCT	GGACATGAGAGGGGAGGACT
82	Lipg-ENSMUSG00000053846-ENSMUST00000066532	CAAGAAGTCCTGGCGTCCAC	ACAAAGCTTCAAGCCAAATGGA
83	Rrbp1-ENSMUSG00000027422-ENSMUST00000016072	TGAGTCTCCTCGGCAAACTG	ATGGCATTGACATCCAGGCA
84	Myo1e-ENSMUSG00000032220-ENSMUST00000034745	GAAACTGACACGTGGGCAGA	GAAGTAAGGGAGGACGGCAG
85	Tmed3-ENSMUSG00000032353-ENSMUST00000058488	CTCCTAGCCTCGAAGTCCCT	CACAGTTCTGCCTCCAGGAG
86	Timp2-ENSMUSG00000017466-ENSMUST00000017610	GGACCCGTAAGAAGGCTGAC	TGAGTTATGCAGCAAGCCCA
87	Icam1-ENSMUSG00000037405-ENSMUST00000086399	GAGCCTGCTGGATGAGACTC	CAGCTTGCACGACCCTTCTA
88	Plod3-ENSMUSG00000004846-ENSMUST00000004968	CCCCTGACACTCAACCAGTC	ATTTGTGCATGGTGGGAGGC
89	Actn4-ENSMUSG00000054808-ENSMUST00000068045	CTTGGCCACCCTGAACAGAG	TAAACTCAAGGGGGCGAGTG
90	Flna-ENSMUSG00000031328-ENSMUST00000114299	TGAGCCTGCCACCTAAGTTG	TCACAGGACAACGGAAGCAA
91	Adgre5-ENSMUSG00000002885-ENSMUST00000075843	GGATGTGAAGGCGAGTTCCA	TGTGTGTCAACATTGAAAGGTTT
92	Itgb3-ENSMUSG00000020689-ENSMUST00000021028	ATGTGTGTGACGAGTGTGGG	GAACAGAGCTACCTGCCAGG
93	Kank2-ENSMUSG00000032194-ENSMUST00000034717	TACAGGGAGGAGCCAGTGAG	GCCCCGACAGAGAAACAGAA
94	Lmna-ENSMUSG00000028063-ENSMUST00000029699	GTAATCTGGGACCTGCCAGG	AGGTGCAGATGGGAAATGGG
95	Gns-ENSMUSG00000034707-ENSMUST00000040344	GGCTGACCCTGTTTCTCCTC	CAGTGAACAGGGTCCGTTGA
96	Cdh5-ENSMUSG00000031871-ENSMUST00000034339	TAGGGTTCTGGTCTTTGGGC	GTACATCTCATGCACCAGGGT
97	Tax1bp3-ENSMUSG00000040158-ENSMUST00000040687	AGTCAGCCACCGTCTAGACT	GCAGGGGACAAGTTAGCCAT
98	Podxl-ENSMUSG00000025608-ENSMUST00000026698	CTCTGATCTGTCTGCTGGCC	AGAAAGGCCCCCACAAAACA
99	Adamts15-ENSMUSG00000033453-ENSMUST00000065112	GGGGAGACGCTTCAACAGAA	TTAGGAGGAGGTCGCTGTCA
100	Cnn2-ENSMUSG00000004665-ENSMUST00000004784	CACCCAGCATGCCAGCTC	GATGTGGGACCTCAGGCAAA
101	Pdlim7-ENSMUSG00000021493-ENSMUST00000155098	GAGATTGCCACCCTACTCCG	CACACTGTTGCTAATCTGGCA
102	Cemip-ENSMUSG00000052353-ENSMUST00000064174	AGTGTTGGCCGATGCTACAA	CTGATGTCCTTGGCTCTGGG
103	Psd4-ENSMUSG00000026979-ENSMUST00000102942	TTTGTCCCCCATGCACTCTC	AGCTTGGTGGAAGACTGCTC
104	Fosl2-ENSMUSG00000029135-ENSMUST00000031017	CTTAGAAGCCCTGGGACACG	CGAGAACAAACCCGAGGGAA
105	Vangl1-ENSMUSG00000027860-ENSMUST00000159388	GACTTCAGCCTCGTGGTCAA	TGGTTCGTGGAATGCACAGA
106	Gmppb-ENSMUSG00000070284-ENSMUST00000047947	GAAGGATATGGAGGGCTGGC	ATTATTCCAAACAGCCCTGGG
107	Cnn1-ENSMUSG00000001349-ENSMUST00000001384	GGCTGGCTGCTGCTCTTG	GATGATGGTAGTCCCGGCAC

108	Cgnl1-ENSMUSG00000032232-ENSMUST00000072899	GACAAAGCCCCATTGTGCAC	TCCCATGGAGGTAGCTGTGA
109	Pttg1ip-ENSMUSG00000009291-ENSMUST00000009435	AGTTCTAAGGTGGCTGGCAC	GTTCTGTGGCTTCATGCTGC
110	Slc7a1-ENSMUSG00000041313-ENSMUST00000048116	GACTATCCCAGCGAGGAAGC	GAGGGGTCTTGCCAGTGTAC
111	Ppfibp2-ENSMUSG00000036528-ENSMUST00000098134	CTGTCCCTGATGCTGGTGAC	TGCTGAAGGGCAAAGTGGAT
112	Copz2-ENSMUSG00000018672-ENSMUST00000018816	TTGAAATGAAGACCTTGGAATCAA	AGGCCTAGGACTGAGGATGT
113	Col5a1-ENSMUSG00000026837-ENSMUST00000028280	CTAGGCTAGGAGCTGCTGAG	TGTACCGTTTGGCCTGGAAT
114	Epha2-ENSMUSG00000006445-ENSMUST00000006614	GCCGTGCCCAACAATACTTG	TCAGTTTGAACCATCTGCAACT
115	Chst14-ENSMUSG00000074916-ENSMUST00000110837	AATGACAGTAGGCCAGCACC	TGCATTAGACAAGTGACAGCCA
116	Slco3a1-ENSMUSG00000025790-ENSMUST00000098371	GCTGAGGGCTGCTCTCTTAA	ACAGACTACCACACGCAGAA
117	Htra1-ENSMUSG00000006205-ENSMUST00000006367	GCAGGAGCCAGACTTCATGT	AGTCACTGTACTCCGGGTTC
118	Fbln1-ENSMUSG00000006369-ENSMUST00000057410	CTGGTTCTGAGGGCAGGTTG	TGACATCCTGCCTCGTCATG
119	Eml3-ENSMUSG00000071647-ENSMUST00000096241	GCTGGATGGTCTTGTCCCTC	CACCGAGCAGAGGGTGTC
120	Ccdc102a-ENSMUSG00000063605-ENSMUST00000077955	TGGCCTAGGACATCCAGGG	AAGTCTGAGTTCATGTTCATGGT
121	Krt13-ENSMUSG00000044041-ENSMUST00000007275	ATTAACTCAGCCTGGCGTCC	TGTAGAGACACTGAGAAGCCA
122	Tmc6-ENSMUSG00000025572-ENSMUST00000026659	TCACCATGATCCTGCAAGCA	TGCACACCTCTGGTTTCAGC
123	Loxl2-ENSMUSG00000034205-ENSMUST00000100420	ACAGTAAAGAAGATCCTGGGCC	GGAGTTCAGCCCGATCACC
124	Wdr1-ENSMUSG00000005103-ENSMUST00000005234	CTCACCCCAAGGACCGAATC	TGCAAGGGGAGGGATTGTTC
125	Mydgf-ENSMUSG00000019579-ENSMUST00000019723	TCAAGGGCCTTCATGTCCAC	ATGGTTGTGAGCCACCATGT
126	Fhl3-ENSMUSG00000032643-ENSMUST00000038684	GGCCCCTGAGTGGAGGAT	AGGGCAACATGGAGCTGAAA
127	Colgalt1-ENSMUSG00000034807-ENSMUST00000047903	CAAACAGCCAAAGCCACTGC	GGCAAGCAGTCATCCCAAGA
128	Serpine1-ENSMUSG00000037411-ENSMUST00000077523	TTGACAGTGGGAAGAGACGC	CCCTATGGGTACACGGTGTG
129	Cd248-ENSMUSG00000056481-ENSMUST00000070630	TGTGTGATGGGGTGCAGATG	TGAGGTTGGCACCCCATTTA
130	Lamc1-ENSMUSG00000026478-ENSMUST00000027752	TAGTGGCGAGAGGGCTGTAA	GCGGGGGAGGGATAAAACAA
131	B4galt1-ENSMUSG00000028413-ENSMUST00000030121	TTTACCCACAACCTGGCTCG	ATGCCAAGGTCAGAGACAGC
132	Rassf3-ENSMUSG00000025795-ENSMUST00000026902	CACAAGGGACTCCACAAGCT	CCAGACACAATTTCAGGGGAGA
133	Arpc1b-ENSMUSG00000029622-ENSMUST00000085679	AGACTAGACTGCTGCCCTCA	ATCAGGGCCTAGATGTCCGT
134	Car12-ENSMUSG00000032373-ENSMUST00000071889	CCTGGCAGTCTCAGACATCC	TGGATACTTCCTCAGGGCCA
135	Bag2-ENSMUSG00000042215-ENSMUST00000044691	CCTTCAAATTCACCGGCCCT	TCCTTTGGCCTGATGAACCC

136	Rab13-ENSMUSG00000027935-ENSMUST00000107373	CTTGCCTCCTATTCACCCTGA	AGCAGCTAACAGTACCCAACA
137	Calr-ENSMUSG00000003814-ENSMUST00000003912	TGAGCTGTAGAGGCCACACC	ACCTGATGGCTACTCTCCCA
138	Sdc4-ENSMUSG00000017009-ENSMUST00000017153	TACGCATGAAGCTTCCTCCC	GGCACCAAGGCTAGAAGAGG
139	Glipr2-ENSMUSG00000028480-ENSMUST00000030202	AAGTAGCTCTCGCCAGCAAG	AACAAAGACCAGCCAGCCTT
140	Tex264-ENSMUSG00000040813-ENSMUST00000163441	TGAGGAGTAATGATAAGCCCTCA	ATTCTGGGTTCTGGAAGCTCC
141	Adrb3-ENSMUSG00000031489-ENSMUST00000081438	GTGAAGGGCCGTGAAGATCC	TGAACCCTGGAATTTCCCCC
142	Dock5-ENSMUSG00000044447-ENSMUST00000039135	CTGACTTTGCACCTCCGAGT	GGAGGAGCTGTCAGTAGGGA
143	Anxa2-ENSMUSG00000032231-ENSMUST00000034756	TGAAGGGCTCAGCACAGTG	ACGGTTTATTCACGACACAGT
144	Mbtps1-ENSMUSG00000031835-ENSMUST00000098362	AGCCAGCCACAGAAGCTAAC	ACGGTGAAATCAAGATCCCCA
145	Trim25-ENSMUSG00000000275-ENSMUST00000107896	GGCTGCTAGTTACTTGGGCA	AGGGCAAGGTCTAGATGGGT
146	Col1a1-ENSMUSG00000001506-ENSMUST00000001547	CGTGTAAACTCCCTCCACCC	GTGGTGCTCTGAAACCCTGA
147	Lgmn-ENSMUSG00000021190-ENSMUST00000021607	TGAACAGCTCGCTTCCCAAT	CCATAGGGGCTTGCTTTCCA
148	Lima1-ENSMUSG00000023022-ENSMUST00000073691	TGAAGAATGACCGGCCGC	CACCCCGCCATACACATCAT
149	Sparc-ENSMUSG00000018593-ENSMUST00000018737	TAAGTTCACGCCTCCTGCTG	TGAAATGCTTGGAGGGGAACA
150	Pxdn-ENSMUSG00000020674-ENSMUST00000122328	GCCTTAGTGTTCCCGAGCTT	CCCAATGCAGTGGCTAGTCA
151	Cyr61-ENSMUSG00000028195-ENSMUST00000029846	GTTCCTAGTGTGGGCTGGAC	CCTCCCTCCCCAAAAGCTAC
152	Gria1-ENSMUSG00000020524-ENSMUST00000036315	AGCAGACAGGAAACCCTTGG	GCCCAGGTACCAAAGAGCAA
153	Ptk6-ENSMUSG00000038751-ENSMUST00000016511	CTGGTCTGAGCTCCTTGGAC	TGCTTCACAAGTGCTGGGAT
154	Mobp-ENSMUSG00000032517-ENSMUST00000174193	AGAAGTGACCGAGAAGGGACT	ACGCGACTCAGTGACGAATT
155	Stxbp2-ENSMUSG00000004626-ENSMUST00000160708	TTGCCTCTGCCCTCTCCTAT	AGACACACCAGAAGAGGACG
156	Parva-ENSMUSG00000030770-ENSMUST00000106643	TGTGGAGTGAGACCCTGGG	CCCTACTTCTCAGGCGATGC
157	Myh11-ENSMUSG00000018830-ENSMUST00000090287	GAGCCACCAGGAGAGGAAAC	TGAAAGTGACCATGGGTGCA
158	Ripk1-ENSMUSG00000021408-ENSMUST00000167374	CTCTGGCAGTGGGAAGCAA	TCAGTGTTCTGGTGGCCATC
159	Efemp2-ENSMUSG00000024909-ENSMUST00000165485	ACCTTCTGAAGACCCTCAGG	TTTAAGCAGAACGGTCATCCCT
160	Kit-ENSMUSG00000005672-ENSMUST00000144270	CCAAGTCCAACAGGCTTTGC	TTCAGGGACTCATGGGCTCA
161	Elmo3-ENSMUSG00000014791-ENSMUST00000109375	AGCCGTGACTCAAGTTAGGC	ACTGGCACTTTCTCCAAAGGA
162	Hspb2-ENSMUSG00000038086-ENSMUST00000042790	CTGACTGCCAGCGATAGACC	GGTTTATTCAGCCCCACCCA
163	Plac9a-ENSMUSG00000095304-ENSMUST00000052286	CTGACTTCCAGGGATGGTGG	ACATATGTGTGTGTGTGTATTT

164	Pla2g12b-ENSMUSG00000009646-ENSMUST00000009790	AAGAGAAGGTGGCTCTTCCAG	TGTGAGGTGTCAACTGGCAA
165	Tnfrsf10b-ENSMUSG00000022074-ENSMUST00000022663	GTGCGTTTGAAGTCAGCCTG	CCACTGCTCTTAACCGCTGA
166	Mcc-ENSMUSG00000071856-ENSMUST00000089874	GCACTCTGGCACCAAAGTTG	CGGGCTCTTGAAGGTAGGTG
167	C2cd4a-ENSMUSG00000047990-ENSMUST00000054500	TGCTATGAGCGCGGCTTC	GACTGAAACAAAGCCCTGGC
168	Tagln-ENSMUSG00000032085-ENSMUST00000034590	TTAGAAAGGGAAGGCCAGCC	TTTCTGCACAGCTTTCCAAAA
169	Gucy2g-ENSMUSG00000055523-ENSMUST00000069183	TCTCAAGATGTTTGTAGAGCCCC	AGCAACCATATAGGCTGGGG
170	Bmp8a-ENSMUSG00000032726-ENSMUST00000040496	GCATCCTGCTTCTACTACCTT	CTCCCGAGACAGGCTCTTTC
171	Impa2-ENSMUSG00000024525-ENSMUST00000025403	CACACACAGCTCGAAGGCTA	CGTGGGCTTCGATAAAGGGT
172	Slco2b1-ENSMUSG00000030737-ENSMUST00000032985	CTGGCCTTGCTTGCTCTTTG	AGAGTCTCATCCCCGAGAGA
173	Lpin1-ENSMUSG00000020593-ENSMUST00000111067	GAGAGGCTGCATCTCATCCC	CTGTCTGTCTAGCCAGCACC
174	Rem1-ENSMUSG00000000359-ENSMUST00000000369	CTCTCTCTGGGGTGATGGGA	GTCTGAGACATCTGGGCCAG
175	Sox7-ENSMUSG00000063060-ENSMUST00000079652	GAGCTGGAGGAATGGAGCC	TGCCTGTGATAGTGTTTCACGA
176	Mettl21b-ENSMUSG00000080115-ENSMUST00000116231	GTCTCAGTGGAGGAATGGCA	ATGAGTCCGGACACACACAC
177	Pik3cb-ENSMUSG00000032462-ENSMUST00000136965	GCAGATATGGCCCAAACTGC	TGAAGTCACGGGGCAAGTAC
178	Pacsin2-ENSMUSG00000016664-ENSMUST00000056177	TATCCAGTGACAGCCCATGG	AGGACTTGACAGCTTTTGGTCA
179	Cers2-ENSMUSG00000015714-ENSMUST00000015858	TGACTGAACCATTGTCCTAGCT	TTGGTCTGAAATGGTCCCTCG
180	Crtap-ENSMUSG00000032431-ENSMUST00000084881	GTCCACAGGGGCTAAGGAAC	CAGGGATGAGAAAGCAGACCA
181	Dock8-ENSMUSG00000052085-ENSMUST00000025831	AAAAAGCCACCTTTGCCGGT	AGGTCACTCATGCCCAATGG
182	Serping1-ENSMUSG00000023224-ENSMUST00000023994	CAGGGGTTGAGACAGGCTTG	GGTTTGAACCGTAAAAGCTGCA
183	Bgn-ENSMUSG00000031375-ENSMUST00000169489	AAGAAGTAGAGGCAGTGGTTG	ACCTCATAAGCAGCACTCCC
184	Prrc1-ENSMUSG00000024594-ENSMUST00000025490	CATGTGAGAGCAGTGCGGG	AGTGTGTGGCAGTACAGAGC
185	Trpv1-ENSMUSG00000005952-ENSMUST00000102526	GGACACCATGAAGCAGCTGA	TTAGCACTGCACAGAGTGGG
186	Asb17-ENSMUSG00000038997-ENSMUST00000044089	AGCATGCATCTCTACACTGTCC	CACTTCGGAGGCAGAGACAG
187	Lifr-ENSMUSG00000054263-ENSMUST00000067190	ACCAGGTCACCCTTTGTCAC	AGAACGCCAGATAGCTGCTG
188	Pvr-ENSMUSG00000040511-ENSMUST00000043517	CGGTGCTGGGTAGACAGAAC	TTGCATGTGGGAGGGCTTAG
189	Sord-ENSMUSG00000027227-ENSMUST00000110551	GATCTATGCCCTCAGCCCAC	CTCAGGTCTGGGCCATGATC
190	Stard8-ENSMUSG00000031216-ENSMUST00000036606	GCTGACATTGCCAGGTGTCT	GGGAGTAGCTGTGTGTGCAT
191	Fabp3-ENSMUSG00000028773-ENSMUST00000070532	CTGCTCCGTCACTGACCG	GGTGGCCTTGGTTCTGCTTT

192	Lars2-ENSMUSG00000035202-ENSMUST00000038863	TAGGACCCTCCTTCTCTGCC	TCCCTGGTAGTTCTTTGAGACA
193	Baiap2l1-ENSMUSG00000038859-ENSMUST00000055190	CGATGAAGATGCAGGGGACA	TGGTTCCTGTACTCGTGTGC
194	Fbxo6-ENSMUSG00000055401-ENSMUST00000105706	GTTAGTGGGCAGTCAGGCTC	ATCGAAGCTGCCCTGAGAAG
195	Serpinh1-ENSMUSG00000070436-ENSMUST00000094154	GTCCAAGAGTGGGCGTGG	TCTGCTTTCACACACCTATAACA
196	Hmox1-ENSMUSG00000005413-ENSMUST00000005548	AAATGCAATACTGGCCCCCA	ACACAGAAGTTAGAGACCAAGGT
197	Patj-ENSMUSG00000061859-ENSMUST00000041284	ACTTTTTCACGAGACTGCTGC	TTAACGCCGCCATTCCATCA
198	Irf4-ENSMUSG00000021356-ENSMUST00000021784	GTGACTGTGCCCTGGCTTAT	ATCTCTGAGTCCTGGTCCCC
199	Lama4-ENSMUSG00000019846-ENSMUST00000019992	CTGACACAGCTGCAAGGCT	GAAGCGGCTGGTCTCTGATT
200	Arhgef5-ENSMUSG00000033542-ENSMUST00000031750	TGAGGAGTCTTCTGAGCTCA	GAATGCCACCCTCACATTGC
201	Dnmt3l-ENSMUSG00000000730-ENSMUST00000138785	AAGTCTTTCCTAGAACCAGGGC	TGAAAAACTTCACAACAGGTACACA
202	Txndc12-ENSMUSG00000028567-ENSMUST00000030296	TGTGACATGAATGAGTGTTCTGG	TCTTGTGTTTGTTTTAACCCCAGT
203	Arhgap6-ENSMUSG00000031355-ENSMUST00000139146	ACACCCTGACCTGGAGAGAC	ACCACTGAAAAATTTCACTGCCA
204	P4ha2-ENSMUSG00000018906-ENSMUST00000093107	TGACGTCCTTTTCTCTCCGC	AAGCCACAACAAATAGAAATGTGA
205	Fkbp9-ENSMUSG00000029781-ENSMUST00000031795	AGATCAGTGTGTGCCAAGGAG	CACTGAATGAGTGCACAGCA
206	Mpeg1-ENSMUSG00000046805-ENSMUST00000081035	TCCAGCTTAATTGTCTCCAAAGGA	AAGGCTGAGACACTTGGAGC
207	Lamp1-ENSMUSG00000031447-ENSMUST00000033824	GCCTGTTCTCACATCCCCAA	GCATCGTTACAGGCAATGCA
208	Fndc1-ENSMUSG00000071984-ENSMUST00000097425	GGCCTGCTGCAAGTTCATC	ACAAATAAGATCTGAGGCACTTAGC
209	Ccna2-ENSMUSG00000027715-ENSMUST00000147380	CAGCTCTGCGGTTGCTATGA	TCTTGACAGTTGGCAGCACC
210	Ptprb-ENSMUSG00000020154-ENSMUST00000092167	AAGCGAACGTCAGAGCTGAA	TGGTGAGGGTTTGTGCAAGA
211	Gxylt1-ENSMUSG00000036197-ENSMUST00000049484	ATGAGACAAGCAACGCCTCA	GCTGAACCTCCAGAACGGAA
212	Ggta1-ENSMUSG00000035778-ENSMUST00000102794	TGTCTGACTTCAAATTGTGATGGA	CCCTCAAATTCCCCAGCCTT
213	Gorasp2-ENSMUSG00000014959-ENSMUST00000112201	CAATGTTCAGAACCGACGTGG	ACTGGAGTTTATTGCCGCGA
214	Capn6-ENSMUSG00000067276-ENSMUST00000087316	TCATCAGAGAATCCTGACAGCT	GTAAGGCAGTGTGGCAAAGG
215	Pag1-ENSMUSG00000027508-ENSMUST00000161949	TCCCATGTGTGCTTCATCCC	AACACCCGTGTCAGTTCCTT
216	Pdlim1-ENSMUSG00000055044-ENSMUST00000068439	GAGAGGACCTGCCCTGTCTA	TAAGGGAGCATTCCAGGTGC
217	Pdgfra-ENSMUSG00000029231-ENSMUST00000168162	AAGCCTGTGTGGATGACTGG	TACCACACCACCATGTTGGG
218	Ckap4-ENSMUSG00000046841-ENSMUST00000167671	AAGTGAACTGCGTGTGTGGG	AAGAACACCCCCTCCAAACG
219	P3h2-ENSMUSG00000038168-ENSMUST00000039990	TGAGCTATAGAAAAGAGAGTGCTC	GACAGCAGCAGGAAGAGGAG

220	Myzap-ENSMUSG00000041361-ENSMUST00000093823	GCACTCAGAGACATACGCCT	TCAAAGTAGAGCGTGCGTGA
221	Tgfbr3-ENSMUSG00000029287-ENSMUST00000031224	CTAGGTGGACAGATGGACGC	GGTTACAGCAGGGGTGTTCA
222	Col12a1-ENSMUSG00000032332-ENSMUST00000071750	TCTAAGTCTCCAGTGCTGCT	TCTTGCAAACATTTCCAACTGCT
223	Ankrd1-ENSMUSG00000024803-ENSMUST00000025718	TCTGAGAAAAGAGACTCAACAGGA	CAGCGCGTGTCCTTTTGATT
224	Thbd-ENSMUSG00000074743-ENSMUST00000099270	GGGATTTGCTCCAGAGACCC	GGCTTTGCTTGTGTTCTAGCG
225	Dennd2c-ENSMUSG00000007379-ENSMUST00000172288	GCCAACGACAGTGTTTTCCC	TCCCAAAAGCAGTCTCCCTG
226	Gadl1-ENSMUSG00000056880-ENSMUST00000069651	TTCTTACCCCCAATGGCACC	GAACAGATAGCTCCCTCCGC
227	Col4a1-ENSMUSG00000031502-ENSMUST00000033898	ACTATGATGCTCGCCTCTGC	TGCAGCAAAGCTTACAGGTT
228	Pdcd6ip-ENSMUSG00000032504-ENSMUST00000035086	CGTGCTGCTGGTTCAGATCA	TGCAAAGGTTTTGTCCAGCT
229	P4ha1-ENSMUSG00000019916-ENSMUST00000009789	TTCCCTTGGCTCCTGTTGTC	GTGTCTCATCTCGGCTCAGG
230	Grem2-ENSMUSG00000050069-ENSMUST00000055294	CAACTGTCCCAGATCCACCC	TGACAACACTGTGAGGCTGA
231	Fabp3-ps1-ENSMUSG00000056366-ENSMUST00000070435	CTGCTCCGTCACTGACCG	GGCCTTGGTTCTGCTTTATTGT
232	Plat-ENSMUSG00000031538-ENSMUST00000033941	AAGCCCAGCTCCTTCAATCC	TCATACAGTTCTCCCAACCATCT
233	Ankdd1b-ENSMUSG00000047117-ENSMUST00000055607	CCCTCCACTGCTCCAGTCTA	CATCCTCCTCCTCCACCCC
234	Bhlhe40-ENSMUSG00000030103-ENSMUST00000032194	GATCTCCTGCTGCCTTGCTT	AATGCCAGGCACATGACAAG
235	Tor1aip2-ENSMUSG00000050565-ENSMUST00000111757	CAGAGGTGACATGGAAGGGC	ACTCTCGTGTTCACAGGCTG
236	Sumf1-ENSMUSG00000030101-ENSMUST00000032191	CTGACAGCCAAGAGGAGCTT	TCACACACCCTGAATGGAGTC
237	Kank1-ENSMUSG00000032702-ENSMUST00000049400	AAGCATGGCCCTTGTGTGTA	GGACCAAAGTAAACGTGGCG
238	Wisp1-ENSMUSG00000005124-ENSMUST00000005255	GGTGTGTGGCTCAGGGTAAA	ATTCAGGGCTGTCACTGTGG
239	Sdf2l1-ENSMUSG00000022769-ENSMUST00000023453	GATGAGTGATGGGAGGGTGG	ATTGAACAGGTTGGGGCCAG
240	Zfp568-ENSMUSG00000074221-ENSMUST00000148442	GAGTGAATGAGTTGCCATCAGA	TGCGCTACCACACCTATTGG
241	Echdc2-ENSMUSG00000028601-ENSMUST00000116309	GGCAAGTGACCCAGCTAACT	TGCAGCATGTTTATTAACACAGGG
242	Vamp3-ENSMUSG00000028955-ENSMUST00000030797	CCGATGAAACTGAAGCCCGA	GTCAGACCAGGACAGAAGCC
243	Slc5a2-ENSMUSG00000030781-ENSMUST00000118169	AGGGTGTTGGATGCCATGAG	GGGCAGTTTTATTTTTGTATAGGCA
244	Lpar3-ENSMUSG00000036832-ENSMUST00000039164	CTAAGCCACGGACGCCTC	AGGGATAACAGGTCCTCGCA
245	Tnc-ENSMUSG00000028364-ENSMUST00000107372	GCGGGCATAAATTCCAGGGA	CTGTCCAGACACGTTTGGGA
246	Cubn-ENSMUSG00000026726-ENSMUST00000091436	GGAGAGCTGTGTGTAACCCC	ACCCGTACTTTGACTTTTGTATAGT

247	Hdc-ENSMUSG00000027360-ENSMUST00000028838	TGGTGTAGAGTCCTCAATCAGA	TGTTTTTCACATTACAAGCTCAACA
248	Rcn1-ENSMUSG00000005973-ENSMUST00000006128	AGACACTGGTTCTGGTGTGG	TGCTCTGTACACAAGCCCAG
249	Thbs1-ENSMUSG00000040152-ENSMUST00000039559	GCCAATCATAACCAGCGCTG	GCACGTTCCTAGGAAAAAGGG
250	Plaur-ENSMUSG00000046223-ENSMUST00000002284	TGGACCTGACTCCTGAGCC	CTGAGGAGGTAGAGGCAGGT
251	Palld-ENSMUSG00000058056-ENSMUST00000121785	CGCCCTTGTTGAAGCTGAAA	AGTTAATCCCAAGCCCAGCA
252	Edn1-ENSMUSG00000021367-ENSMUST00000021796	TTGACTACAGAGCTCCCCCA	TAGACCAATATGGCCTGCCC
253	Hacd4-ENSMUSG00000028497-ENSMUST00000151280	AACGTGTGAGGTAGCCTGTC	GTGACAGTCACAGGACCCAG
254	Agtr1a-ENSMUSG00000049115-ENSMUST00000066412	GGTTCAAAGCTTGCTGGCAA	GCCAAAACTTCATCAAATGACAGC
255	Dram1-ENSMUSG00000020057-ENSMUST00000020249	AAGGTATGTTTCCCTGGAGGA	TCTCGGGTACAAAGTCCAAGG
256	Calu-ENSMUSG00000029767-ENSMUST00000031779	TGCAGACAGAGGAACCTACA	TGCTCACTGCCTCAGACTTG
257	Mertk-ENSMUSG00000014361-ENSMUST00000014505	AGCTGAGAGGAGGCATGAGA	CTCCTTGGCTCTTTCCACCA
258	Csrp2-ENSMUSG00000020186-ENSMUST00000020403	TGAACCCAGTAAGCACGACA	TTCACTGGAGGAAGGAACCA
259	Gata6-ENSMUSG00000005836-ENSMUST00000047762	GGTGCTACCAAGAGGCAAGG	TGTGGGTTGGTCACGTGGTA
260	Itgb1-ENSMUSG00000025809-ENSMUST00000090006	TCAGGCGGATTTTGCAACAC	TCGAGACAGAGCAAGCATGT
261	Pdia3-ENSMUSG00000027248-ENSMUST00000028683	AAGCAACAGCCAAATGCACC	GGTTCTTGTTCTTGCCAGCG
262	Rpia-ENSMUSG00000053604-ENSMUST00000066134	TCTAGCCCTGAGAGGAGCAG	CAAGGGTCATCGGAGGAACA
263	Plac9b-ENSMUSG00000072674-ENSMUST00000100809	TGACTTCCAGGGATGGTGGA	CGGAGCTTTTATTTCACACAAAGG
264	Myl12a-ENSMUSG00000024048-ENSMUST00000148960	AGATGACTGAAGAGCTGCGG	ACCTCACTTTCCCAGGCATG
265	Prss35-ENSMUSG00000033491-ENSMUST00000036426	AGGGAAAACTGAAGCGAGCT	GCCACACGTTTTGAGGACCA
266	Pdia6-ENSMUSG00000020571-ENSMUST00000057288	GCCTTCCTTGGGCTTTCCT	TCAGTCCTGAGTCTTTTCTTGTGT
267	Vcl-ENSMUSG00000021823-ENSMUST00000022369	CCTCGTCAAATCTGGCTGGT	TGGTGGACATGTGTCTCCAG
268	Lyn-ENSMUSG00000042228-ENSMUST00000041377	TGGGCATGGAGACTAGCCTA	GTAGGGACTGGAGGGCCTTA
269	Tpm4-ENSMUSG00000031799-ENSMUST00000003575	ACCAGAAGAGTACTGTTTTAACACC	TGCTCAAAGTTCATCTCCCTGG
270	Klb-ENSMUSG00000029195-ENSMUST00000031096	TCAGCTAAACTGCCATTTCTGTC	AGAGAAAAAGGATGTGAAGTCGTG
271	Bag3-ENSMUSG00000030847-ENSMUST00000033136	CTCCCTAGTGTCCCCTGGG	ACAACATTAAAGCCATTCCAACA
272	Slc16a9-ENSMUSG00000037762-ENSMUST00000046807	TGTTTAGAGGGATATTAAAAGGCCT	TCCCCTGGAGCTGGAGTTAG
273	Piezo2-ENSMUSG00000041482-ENSMUST00000047480	TCACTTAAGACCTGGACTGCA	AGTCAGTCCTGAACCCAGAGT

274	9130008F23Rik-ENSMUSG00000054951-ENSMUST00000068258	AGTGACACAGCTTATGGGGC	TGCTGTATGGCACTACTCCT
275	Mvp-ENSMUSG00000030681-ENSMUST00000165096	ACAAAAGTGACCTGGCGGC	TCCATTGTCAGTGTAGTCAGTGT
276	Lamb1-ENSMUSG00000002900-ENSMUST00000169088	ACAGGAAGGGGCTGTAGAGG	CCCAGCAACTGAAGGGTTCT
277	A730049H05Rik-ENSMUSG00000048636-ENSMUST00000057977	GGTTTGCAAAGGCCAAGGAC	CCAACACTGTGGGAACAGAA
278	Lrp2-ENSMUSG00000027070-ENSMUST00000080953	GCCTTGCCAACTCTCTAGGG	TCCAGCAGCATCATTTCCAGT
279	Cat-ENSMUSG00000027187-ENSMUST00000028610	CCTGTAACTCCGGTGCTCAG	CAGACACCATGGAGGCTCTG
280	Kcnq5-ENSMUSG00000028033-ENSMUST00000115300	ATGCATGTTTAGGGGTGGCA	AGCTTGAAATGGGTGTGCTA
281	Man1a2-ENSMUSG00000008763-ENSMUST00000008907	CACAGTGGAGCAGCTCTCAA	AGCGAATTTGAGGCCAGTGA
282	Fndc3b-ENSMUSG00000039286-ENSMUST00000046157	TAAATCCAGCAGGCCAGTGG	AGCCTGGAATGTATGTCCAACA
283	Cd28-ENSMUSG00000026012-ENSMUST00000027165	TGACAGGGACCCCTATCCAG	GGCACTGGGGATACAGCAAT
284	Lamb2-ENSMUSG00000052911-ENSMUST00000065014	GACCACTCCCTAGGGCCTAG	GGAGGTTCACACAGGTTTATTGA
285	Heatr5a-ENSMUSG00000035181-ENSMUST00000040583	CTCTCCTAGCACACTGAGCG	CACCACGTGCTCTTGGTCTA
286	Ctgf-ENSMUSG00000019997-ENSMUST00000020171	GCCAGGAAGTAAGGGACACG	CCCACCCCAAACCAGTCATA
287	Zc3hav1-ENSMUSG00000029826-ENSMUST00000114900	GGATGTATACCATGCTGAAACCA	AAACTGGGGGCTTGAGAGTC
288	Ptger3-ENSMUSG00000040016-ENSMUST00000173533	AGATGATGAACAACCTGAAGTGGA	CCCCCTTTTTATTGTGCTGCC
289	Actn1-ENSMUSG00000015143-ENSMUST00000021554	GACCTCTAACCCGCCCCT	ACAGTCTGTGCCAAGCAACA
290	Papss2-ENSMUSG00000024899-ENSMUST00000025833	CTGGCTCTGGCTTCTTCCTC	TACCAGCGGGACACTGAAAC
291	Lama1-ENSMUSG00000032796-ENSMUST00000035471	TGAGCCCTAAACTGTCGCC	TAATAGAGGTATCTGTTGTCCGTGT
292	Slc16a1-ENSMUSG00000032902-ENSMUST00000046212	CTGACCTGTGAAGCCTGGAG	TGGACACGTTCCTGATCAGTG
293	Fam114a1-ENSMUSG00000029185-ENSMUST00000031080	GCAGGCCACATCTCTTCACT	GGTTGTTTCAGGAGACATACAGC
294	Mme-ENSMUSG00000027820-ENSMUST00000029400	AGATACTGAACATCCACGGCA	AGCAGAAGGACAAGGAAGGC
295	Ror1-ENSMUSG00000035305-ENSMUST00000039630	GGAAAGACACCAGCCAGGAA	GCTTGTAGTTGGAATGCCCAC
296	Car8-ENSMUSG00000041261-ENSMUST00000066674	GAGAGACCGGGAACAAACCC	GGCCAGTCTGGTCTACACAG
297	Eogt-ENSMUSG00000035245-ENSMUST00000054344	ATCCTCTTGCCTGAGCCAAC	TCCAAAGCTACACCCAAAATGT
298	Naalad2-ENSMUSG00000043943-ENSMUST00000172171	GGCTTCTGTCGAGTAGCTTCA	GTCCCACCTTCAGAGCTGAC
299	Birc7-ENSMUSG00000038840-ENSMUST00000108875	CCAGGAACTACACTGCAGAGT	AGGGTGAGAACTTCTGCTTTAT

300	Rhox2h-ENSMUSG00000079627-ENSMUST00000115162	TCAGAGGTCCTCTTCCTGCA	TCAGCTCTTTATTGCAAATACAAGT
301	Atp10d-ENSMUSG00000046808-ENSMUST00000126664	GCTAAACATGCAGGACACCC	AGGCAGCTTCGTTTAGACACA
302	Fam3c-ENSMUSG00000029672-ENSMUST00000163963	CAAGACTGACGGGAAGGAGC	TCTGGATTTCTTTGGAAGGTCAA
303	Blvrb-ENSMUSG00000040466-ENSMUST00000037399	CACTCTGACCTAGGTGGGGA	AGCAGTAGAGCGCCTGTCTA
304	Slc38a4-ENSMUSG00000022464-ENSMUST00000023101	CACCACTAATCCCGGGGAGA	GGCCACACAAATGATCGGTG
305	Itga1-ENSMUSG00000042284-ENSMUST00000061673	GTGACCCAGAAAGGCTACCC	CCAAACGGCTCTTCCTCTCC
306	Pcgf5-ENSMUSG00000024805-ENSMUST00000062389	CCCAGGCTGAGCTGAATGAA	TAGGTCAGTGGCTCAGAAACC
307	Dnajc10-ENSMUSG00000027006-ENSMUST00000028392	TGAGCTTTGATGGTGTTGAAGA	TTTTGGAGTTGCAGCAGGGA
308	Srgn-ENSMUSG00000020077-ENSMUST00000160987	CTCTGTCTCCCCACCTCCAT	GTTCCAGATGACATCTCCCCC
309	Enpep-ENSMUSG00000028024-ENSMUST00000029658	AGGACGGACAGTGAAAACACG	AGCTTTGCTGGAGGACAACC
310	Nid1-ENSMUSG00000005397-ENSMUST00000005532	TCCTTGTTCCCTCTCCAAGT	TGTACAGCTGTAAACTTGAGGA
311	Ikbip-ENSMUSG00000019975-ENSMUST00000020149	ACTGCACACCTTCTGATGGG	TTGAAACGCAGGAGCAGGAA
312	Adamts9-ENSMUSG00000030022-ENSMUST00000113438	GGCTCTGAAGAGGAAGCCAT	CAAACACCAAGTGCTCTGGC
313	Myof-ENSMUSG00000048612-ENSMUST00000172095	TAGTGAAAGCAGCGCCTTCA	TACTGGACCCACGGGATCAT
314	Prkaa2-ENSMUSG00000028518-ENSMUST00000030243	GGTAGGGGTGGTTCAATGGG	TAGGGGCTTTGAAACGGCTT
315	Erp44-ENSMUSG00000028343-ENSMUST00000030028	ATACTGCAAGCCTTTCCACA	TGGCACATGTCAATCTTTTAAGCT
316	Ppic-ENSMUSG00000024538-ENSMUST00000025419	TGGTGACAGAAATGGCGGAA	ATCCAAGCCCCAGAATGCTC
317	Hhip-ENSMUSG00000064325-ENSMUST00000079038	CTGCATACACACGCACAACC	GCTCTTTTGCCTCCAACGTG
318	Cpox-ENSMUSG00000022742-ENSMUST00000060077	GCATCAGACAGAGCCTTGCT	GGCAAACACAATCACATACATCCA
319	Tek-ENSMUSG00000006386-ENSMUST00000102798	ACTCTTCATGTACAACGGCCA	AGAGGAACCATTGCAGCAGG
320	Anxa5-ENSMUSG00000027712-ENSMUST00000029266	GATGACTGAGAGCCGCCTG	TGGAGGGAAGGGAATGTTTTT
321	Pdlim5-ENSMUSG00000028273-ENSMUST00000029941	TCAGTGGTTCCAGGAAAGAGA	TTGTGTGAGTTGCCTCTGCT
322	Nxf7-ENSMUSG00000031410-ENSMUST00000113163	TGACATAAAAACACACGGCTGG	TCTCAAGAGAAATACACAGAGCA
323	Col8a1-ENSMUSG00000068196-ENSMUST00000089332	TGTTTAGTCCTGTGCACAGCA	ACTCAAGCAAGTTACCGGTCT
324	Flrt2-ENSMUSG00000047414-ENSMUST00000057324	AGCCTAGAGGTCCAGCGTTA	AGAATCTTCCAGCCCCTTGC
325	Otud4-ENSMUSG00000036990-ENSMUST00000173078	TGGACATGAATTCAAGGGGCA	AGCATCACACCATGTAGGGC
326	Colec12-ENSMUSG00000036103-ENSMUST00000040069	CTGATGCCTCTGAAAGCCGA	AGTACCGCTATTATTATCTTTGGCT
327	Manf-ENSMUSG00000032575-ENSMUST00000159620	TGAACCGCAACCATAAGCCT	TGAGGAAGCTGTGGATATGGAC

328	Dab2-ENSMUSG00000022150-ENSMUST00000080880	TTGCCTAGCTTCTGAAGTTGT	ACAAAACAGTTGTGTGGCTGA
329	Cald1-ENSMUSG00000029761-ENSMUST00000115027	CGGAGAAAGAACCCAACTGC	AGTCAGCTTGCAGTTTCCTCT
330	Nog-ENSMUSG00000048616-ENSMUST00000061728	TTCCTGCTAGAACTCGGGGG	TCCTGTTCTGCACTTCTTCCT
331	Pawr-ENSMUSG00000035873-ENSMUST00000095313	ACAAGGTAGAAGCTGCACGG	AGCTTTCCTTTGAATGATTACTGCT
332	Nid2-ENSMUSG00000021806-ENSMUST00000022340	ACGAAGGCAAAGACAGTGGA	GAACTGCTGCATCTTTGGTGT
333	Anln-ENSMUSG00000036777-ENSMUST00000040912	GAAGCCTTAAGCCGAGGAGC	TGGCAAGTGATAGACTTAGTGTTT
334	Rras2-ENSMUSG00000055723-ENSMUST00000069449	AGAGTCCCTTGAGGTTTAGCT	TCAGGAGCGCACTTCGAATT
335	Snx9-ENSMUSG00000002365-ENSMUST00000002436	ACAAGAGAACTCCGAAGTGCT	TCTGACATTGGGTAAGAACTGGA
336	Phldb2-ENSMUSG00000033149-ENSMUST00000076333	TTTGGCCATGGTCCACTTCA	ACGGTCATGATCTCGGTCCT
337	Tmem45a-ENSMUSG00000022754-ENSMUST00000023435	AGTCTTTCTTGATAAGCCTTCTCC	TCTGAGAACTCCCTGGGTTGA
338	Atp13a3-ENSMUSG00000022533-ENSMUST00000100013	AAGGCTGCTCTCTGTGCTAG	CCACGGTCTCAGGGAATTCA
339	Acta2-ENSMUSG00000035783-ENSMUST00000039631	CCCCCTGCTCTGCCTCTA	AGTCACGCATGCATATCCATG
340	Adgrl4-ENSMUSG00000039167-ENSMUST00000046977	AGAAGTCATGATAATTACAGCTGC	CAGCATGCTGACTTCCCTCA
341	Fut8-ENSMUSG00000021065-ENSMUST00000062804	CTTCAGACCATCTCGGCCAA	AGCTTTATAGACAAGAGCAAGCAC
342	F3-ENSMUSG00000028128-ENSMUST00000029771	CTGAAGCCGCTAACGCTCA	TGTCTCAATTCCCAATCACCTT
343	Gm5082-ENSMUSG00000054258-ENSMUST00000067176	ACGTAGTTCTTGTGGCTTTGC	AAAAGAGCCAGTAGCCGCC
344	Itgav-ENSMUSG00000027087-ENSMUST00000028499	CAGACAGCACTGAGTCAGCA	GGCACCACACGTTCAAGTTT
345	Chn2-ENSMUSG00000004633-ENSMUST00000046856	CACCAGGGAAGTGAGCTGAT	CCATAACCAACGCACTGCAT
346	Hebp1-ENSMUSG00000042770-ENSMUST00000045855	TGAACTGAGAACGTCCTGGA	ACTGCATACGTCAAAGATTTGTAT
347	Anxa3-ENSMUSG00000029484-ENSMUST00000031447	TGATTGAAGAAGATGGCTCCC	GCACCAGGCAACATACAGTA
348	Bmper-ENSMUSG00000031963-ENSMUST00000071982	TCAGCGATGACCTTTGTTCT	ACATCAGTGAGTTGGGGGTG
349	Plod2-ENSMUSG00000032374-ENSMUST00000160359	ACGTAACTCAAACTGAATGGCT	TGGCACCGTGAATAAAGACT
350	Col1a2-ENSMUSG00000029661-ENSMUST00000031668	TCTCTTTGCCGTTTCCTCCT	TCTTTTGGAAAACCAAAAATGCTTT
351	Fibin-ENSMUSG00000074971-ENSMUST00000099626	GGCTAATATTTCCATTGTAACTGCT	GGACCTTTGTCACACATTTCCA
352	Actc1-ENSMUSG00000068614-ENSMUST00000090269	GTGCTTCTGAGATGTCTCTCTCT	AGGTTGCAAGTCCTGGTCTG
353	Gm6880-ENSMUSG00000079536-ENSMUST00000114125	ACCAGATACAAGAAAACTCAGAGC	TTTGTCCTGACCAGTGGAGC
354	Matn3-ENSMUSG00000020583-ENSMUST00000020899	ACTCTCGCCTGGAAATGTGG	AATGTCACAGTGAGGGGCTG
355	Isyna1-ENSMUSG00000019139-ENSMUST00000019283	CACTGCTTAAATCAAGTGACCGG	GCTTGAGACTTGCTTTATTGTCT

356	Msl3-ENSMUSG00000031358-ENSMUST00000033725	TGATGGTTCTGTAAAAACAACTGCA	TCAAACAAAACTGGCAAAGCT
357	Igsf1-ENSMUSG00000031111-ENSMUST00000033442	AATTACCCTCTCCTTTATAAGAGCT	TTTTAACAGAATTGCATGCACCA
358	Sox17-ENSMUSG00000025902-ENSMUST00000027035	TTGCCGACCCGACCTGAG	GCAAATTTTGTGGGAAGTGGGA
359	Ptgs2-ENSMUSG00000032487-ENSMUST00000035065	TGAGCTGTAAAAGTCTACTGACCA	CCCAGGTTCAATCAGCAGGT
360	Lmo7-ENSMUSG00000033060-ENSMUST00000100337	CATTCGAAAGCGCTGTTGCA	AGCGGAACAGAAGAGCTCAC
361	Sec24d-ENSMUSG00000039234-ENSMUST00000047923	CCTCGGCTCTCCCTCTGAG	ACCATCAAAATTGTACAAAGAACCA
362	Krt23-ENSMUSG00000006777-ENSMUST00000006969	AGCCAAGGAGTTTCTGCCTG	GAGAGCAGGAGAGTAGAGCT
363	Nov-ENSMUSG00000037362-ENSMUST00000050027	TAACAGGTGCTCAGGGAAGC	GGCTGAAGACGATGGTCCAA
364	Slc6a15-ENSMUSG00000019894-ENSMUST00000074204	TAGCTGAACGAAAGCAGGGG	TGCCATACAGTTTTAACAATCACT
365	Dse-ENSMUSG00000039497-ENSMUST00000048010	TCAGTGCTAACGCCAAGGC	TCACACCAAAACTTTAGTTCCATGT
366	Mbnl2-ENSMUSG00000022139-ENSMUST00000088419	CCTGAGACCAGCTGTGATGT	TGGTTGTACATTGGGTGAGTGA
367	Fst-ENSMUSG00000021765-ENSMUST00000022287	ACTCTCCACCAATGTTCAGTGT	ACATTGCACATTTACATGGTAGAA
368	Ngf-ENSMUSG00000027859-ENSMUST00000106925	CTGACTTGCCTGCAGCCC	TCGGTATACAGGATGCTTTGAAAA
369	Vnn1-ENSMUSG00000037440-ENSMUST00000041416	CTTGTGCTGATGGAATTTTTACATT	AGGTTTGCGTTCTCCTGAGG
370	Clcn5-ENSMUSG00000004317-ENSMUST00000115746	GAATCATGGGCTTGTTGTGGA	ACCGTAGCCATCTTGTTTACTCT
371	Filip1l-ENSMUSG00000043336-ENSMUST00000159816	TAACTGACCACCCTCACCCT	ACTTCTGATAGTTCTCTTTTGAGCA
372	Gm10110-ENSMUSG00000062093-ENSMUST00000081204	GCCAAATAACCCCTCATGGC	AAGTCTTGGGTACATCAAAACCA
373	Igfbp7-ENSMUSG00000036256-ENSMUST00000163898	AACCTGCGAATCCATGAGCC	TTGTTTTGTTTAAAGCTACATGTGC
374	Serpinb6a-ENSMUSG00000060147-ENSMUST00000076532	CTTTGACCCTCTCTCCATGGT	CTGAAGGTGGGGTGCCATAA
375	Serpinb6a-ENSMUSG00000060147-ENSMUST00000167163	CTTTGACCCTCTCTCCATGGT	CTGAAGGTGGGGTGCCATAA
376	Col3a1-ENSMUSG00000026043-ENSMUST00000087883	ACTCTCTGAAACCCCAGCAA	AAGATCAACAACTACATACAAGCTT
377	Nppb-ENSMUSG00000029019-ENSMUST00000103231	GAAGACCTCCTGGCTGCAG	CCTTTTCAAAAGCAGGAAATACGC
378	St3gal6-ENSMUSG00000022747-ENSMUST00000114357	TGACCCTATGGATCCAAAAGATGA	CCAGTCACATGCCCATGAGT
379	Hs3st1-ENSMUSG00000051022-ENSMUST00000117944	GCACTGATTTGCCGTCTCCT	AGATTTGACAAGGAGGTGAAAACA
380	Trpc6-ENSMUSG00000031997-ENSMUST00000050433	TAGAGCAGAGCCCCTCAGAA	CAGCAATTGTTAGCCTCAGCA
381	Plscr2-ENSMUSG00000032372-ENSMUST00000034932	ACAGAAATCCGACCTGCAGT	AGAATTTCCCAGGCAATAGTATCT
382	Cct6b-ENSMUSG00000020698-ENSMUST00000100722	TGAATTCCAAATCAATTCATCTGGA	AGACCAAAAGACATGCACTACCT

383	Icam4-ENSMUSG00000001014-ENSMUST00000001040	TATAGATCTGTTTTCGATGCCTGAC	CTGAGGAGGAAGACTTTATTGTTTC
384	Gsta2-ENSMUSG00000057933-ENSMUST00000034902	CAAGTTTTAGTGTGGCTGCAT	ATTATTGCACAATAGCCAGAATCA
385	Arc-ENSMUSG00000022602-ENSMUST00000110009	AGCCTGAATAGAGGGGCCA	GGGCCTACCTACACACCCTA
386	Arpc2-ENSMUSG00000006304-ENSMUST00000113820	CCCGCTAACTCTTGGGAACT	TGACCCCAACCGTATTTCTTT
387	Arpc4-ENSMUSG00000079426-ENSMUST00000156898	ATCTAGCTGGATCTCGCAGC	ACCTCTCCTGTAAGCCCCAT
388	Camk2a-ENSMUSG00000024617-ENSMUST00000025519	CATTGAAGGACCAGGCCAGG	AGACAGGAGATCCCCGTGAA
389	Cdipt-ENSMUSG00000030682-ENSMUST00000032920	TGATCCTCTCCCATCCCCAG	AGGACAGGAGAGAGGAGCTG
390	Dlg4-ENSMUSG00000020886-ENSMUST00000108589	GACTCTGATTCCTGCCCTGG	CATTTCCTGTCCTCCCCTGG
391	Dlgap4-ENSMUSG00000061689-ENSMUST00000169464	GCTCTAAGGCCGTGAAGGAG	GGCACACAGAGGTCAAGTGA
392	Eif3a-ENSMUSG00000024991-ENSMUST00000025955	ACGACGTTAAGTCCCAAGATGA	CACAGCAGCAAATTCCGTGT
393	Eif3f-ENSMUSG00000031029-ENSMUST00000033342	ACCTGTGAATGAGCCCCAAG	AGTTAGAAACAGCCTGCCCA
394	Eif4a3-ENSMUSG00000025580-ENSMUST00000026667	ACCTCATCTGAAGCTGGTGC	TGTCCACGTTTAATGTAAAAGGTT
395	Eif4e2-ENSMUSG00000026254-ENSMUST00000113235	GGCTCAGGCAGTTCTTCTGT	TAAATCTGCCAGTGCTGCCA
396	Eif4ebp1-ENSMUSG00000031490-ENSMUST00000033880	ACATTTAAGGGACCAGCCGT	CCAGGCTCAGAGGAAAGGTG
397	Eif4g1-ENSMUSG00000045983-ENSMUST00000044783	ATCACAACTGATGGCGGGTG	CCCCTTTCCCCTGCCCAG
398	Eif4h-ENSMUSG00000040731-ENSMUST00000036125	TGAGCTTAGGTTGGGAGGGA	GCCTGTCTTTTCCGCATTGG
399	Eif5a-ENSMUSG00000078812-ENSMUST00000043419	AATAACCGGCTTCCAGGGTG	GAGAGAGCCCACAGGAAAGG
400	Epha2-ENSMUSG00000006445-ENSMUST00000006614	CTATCTGAGTCCATTGGGGCC	TCAGTTTGAACCATCTGCAACT
401	Fxr2-ENSMUSG00000018765-ENSMUST00000018909	AAAACCTCCGGTCCACATCC	TACGCCTTCTCATCCACCAG
402	Gabrd-ENSMUSG00000029054-ENSMUST00000030925	AGTGTGTGCACCAGGTGATG	TCATGCTGGGACCAAAGAGG
403	Gria1-ENSMUSG00000020524-ENSMUST00000036315	AGCAGACAGGAAACCCTTGG	GCCCAGGTACCAAAGAGCAA
404	Htr2c-ENSMUSG00000041380-ENSMUST00000096299	GTGTAAGCAATAGCAGCGCA	TAAGGCAGTCTGTTGCACGT
405	Htr5b-ENSMUSG00000050534-ENSMUST00000055884	AGAAGGGCTGGGGAGAGAAA	TTGAAGACACCTGCGTGGAA
406	Inpp5d-ENSMUSG00000026288-ENSMUST00000072999	CTCCCCAAAGCCCAAGAGG	CAGACTGGGAAACACAGGGG
407	Kcng1-ENSMUSG00000074575-ENSMUST00000109191	GTCCTCCCTCTAAGCCCAGA	GGCTCTATCGCTTCTGGTCC
408	Kcnk12-ENSMUSG00000050138-ENSMUST00000055221	CCTCCAGGTAGAGCGGCC	AAGAAAGCGCCATCTCGGAG
409	Kcnk5-ENSMUSG00000023243-ENSMUST00000024011	GCTCCCCACTTGCTTTCCA	ACATCTGCCTTCACCATCGG
410	Kcnq5-ENSMUSG00000028033-ENSMUST00000115300	ATGCATGTTTAGGGGTGGCA	AGCTTGAAATGGGTGTGCTA

411	Map2k3-ENSMUSG00000018932-ENSMUST00000019076	AGGATTCATAGGGACTGGCC	ACAGCAATAGGTTCCTGCCC
412	Map3k1-ENSMUSG00000021754-ENSMUST00000109267	TGTTCAGATCAGCTCTAATGGAGA	GCCCTTAAGCCCTGGAAGTC
413	Mtap-ENSMUSG00000062937-ENSMUST00000058030	TGGCTGACACAAGAAGACGT	CCCACACGTCCATCAACACA
414	Pip4k2c-ENSMUSG00000025417-ENSMUST00000013970	CTGACTCCAGGGTGATTGCG	AAGGTTGCTTGAGGGATGGG
415	Pip5k1b-ENSMUSG00000024867-ENSMUST00000025800	ATAAGTGAAAGCGGCAGCCA	AGGCCTTTTATAAATAGCTTTGCTT
416	Ppm1m-ENSMUSG00000020253-ENSMUST00000140761	CCACTGAGGGCTCACATACTG	TAAGCAAATGTGAGACTGCCCA
417	Ppm1n-ENSMUSG00000030402-ENSMUST00000032560	AGGAGTGGTGAGATGGGGC	TCACGTCTATCACCACCTCCT
418	Ppp1ca-ENSMUSG00000040385-ENSMUST00000046094	AATAGCCTCCATGTGCTGCC	CCTTTATTCAAGAGACCAGATGGG
419	Ppp1r16b-ENSMUSG00000037754-ENSMUST00000052927	GAACTGGAGGAGAGAGCTGC	CTAGGTATGCAGAGCCAGGC
420	Psd4-ENSMUSG00000026979-ENSMUST00000102942	TTTGTCCCCCATGCACTCTC	AGCTTGGTGGAAGACTGCTC
421	Ryr2-ENSMUSG00000021313-ENSMUST00000021750	ACCTTGATTGTCTCTGAAAACCA	TTCGCAGAATGACCTCGCTT
422	Shank1-ENSMUSG00000038738-ENSMUST00000107938	GAAGAAGGAGGGAAAGGCCG	CAGGTCCTATGCTGAAGGGC
423	Sipa1l1-ENSMUSG00000042700-ENSMUST00000166429	CACCTGCGGGAGGACAAC	AATAGCTTGCCCCAAGTCCC
424	Slc17a5-ENSMUSG00000049624-ENSMUST00000052441	ACAGAAACTGAAGACAACAAACA	AAGAAATGTCACATGGCAACTT
425	Slc17a7-ENSMUSG00000070570-ENSMUST00000085374	GACTACTGACCACGGGCCTC	AGCGTTTATTGGGAGTGGGG
426	Slc17a9-ENSMUSG00000023393-ENSMUST00000094218	CTGACGATCCAGTCACCCAG	AAGGCTCTGGCAAGACTCAC
427	Slc30a2-ENSMUSG00000028836-ENSMUST00000105874	GAATGACTCCCCAGCCAGAC	TTACTGTGCCAGGACAGAGC
428	Slc30a4-ENSMUSG00000005802-ENSMUST00000005952	AGAATGAGCTGTCACCCTGC	ACTTGGTTTTTGGAGGGGCA
429	Slc30a5-ENSMUSG00000021629-ENSMUST00000067246	GCATGAACAGTGAACAGCGG	TCAGTGGTTGTAATTTCTCCTGA
430	Slc30a7-ENSMUSG00000054414-ENSMUST00000067485	CTGACTGCTGAGGGTTCTGG	AGGGAAGTGAAGTGCTGAGC
431	Slc4a10-ENSMUSG00000026904-ENSMUST00000112480	AGAAGCTGATTCCCCAAAGCA	TGACCAGGAATCTGAAATTAGGAGT

RESA barcoding oligos

name	sequence
FW1	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
FW2	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
FW3	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
FW4	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
FW5	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
REV	AATGATACGGCGACCACCGAGATC

4.1.6. Plasmids

piggyBAC transfer vector

piggyBAC transposon vector

4.2. Methods

4.2.1. Cell culture

4.2.1.1. Culture and differentiation of N1E-115

N1E-115 cells were obtained from ATCC. Cells are cultured in Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10% Fetal Bovine Serum (FBS) and 2mM L-Glutamine at 37°C and 5% CO_2 on untreated cell culture plastic. Cells are splitted a least twice per week, one to five, and maximally cultured to a passage number of 30. For splitting, the growth medium is deprived and the attached cells are washed once with prewarmed PBS solution. They are then treated for 5min with 1ml/25cm² PUCK's saline solution at 37°C and 5% CO_2, which leads to detachment. The detached cells are washed off the dish with 3ml/25cm² prewarmed growth medium and singularized by pipetting against the wall of the dish. Afterwards they are spun down for 5min at 800g and the supernatant is disposed of. The cells are resuspended via pipetting in 10ml of prewarmed growth medium and 1/5 of this suspension is transferred to a new culture vessel, containing already warm growth medium. For storage purposes the cells are frozen in growth medium, supplemented with 10% DMSO at -160°C. To do so the cells are washed with prewarmed PBS, detached using PUCK's saline solution, spun down in growth medium and resuspended in freezing medium. Freezing cultures are prepared with a cell density of 2*106 cells/ml in 1ml. For thawing frozen cultures, the cell vial is quickly heated to 37°C in a water bath. After thawing, the cells are spun down for 5min at 800g to remove the DMSO. Afterwards the cells are resuspended in prewarmed growth medium.

For differentiation, the 50% to 70% confluent attached cell culture is deprived of growth medium and washed with prewarmed PBS. Then the cells are cultured in differentiation

medium (neurobasal medium with 2mM glutamine and no FBS) for 2 days to reach the final differentiation state.

4.2.1.2. Culture and differentiation of Ascl1-mESC

Ascl1-mESC were created in the lab by Alessandra Zappulo, following the protocol of Iacovino et al. [283]. The cells are cultured in a mixture of 20% mESC and 80% 2i medium (20/80). Cells are split one to three every day or one to five every other day to a maximum passage number of 25. If the cells are not split, the medium has to be changed every day. Cell culture plastic has to be coated with a 0.1% gelatin solution using a treatment of 5min with 3ml of coating solution for 25cm² of surface. For sub culturing, the cells are deprived from medium and washed with prewarmed PBS. Afterwards they are treated with 1ml/25cm² cold TripLE for maximum 3min at room temperature or until they detach from the plastic. Detached cells are collected in 3ml/25cm² prewarmed Feeder's medium and centrifuged for 5min at 160g. The supernatant is disposed of, and the remaining cells are resuspended via pipetting in prewarmed growth medium. An appropriate fraction of the cell suspension is transferred to a freshly coated culture vessel containing prewarmed growth medium. For storage, the cells are frozen at -160°C in 50% ES-FBS, 40% growth medium and 10% DMSO in a concentration of at least $2*10^6$ cells per ml. To obtain uniform differentiation, the cells are detached from the cell culture plastic using TripLE, washed, singularized and centrifuged in Feeder's medium and cultured for two days in AK medium in uncoated cell culture plastic. The concentration of cells should be around $1*10^6$ cells in 10ml for a 10cm petri dish. These cultures are suspension cultures and the cells will aggregate to form embryoid bodies after few hours. After two days of pre-differentiation, the embryoid bodies are collected via 3min centrifugation at 140g and split carefully into two 10cm petri dishes, each containing 10ml of AK supplemented 3µg/ml doxycycline. This second step of differentiation also lasts two days, after which the cells are collected once more via centrifugation and plated for the final step of differentiation on a Poly-L-Ornithine coated surface in monolayer medium, supplemented with 3µg/ml doxycycline. The coating is critical for the attachment of the cells and is done using a final concentration of 0.1mg/ml Poly-L-Ornithine in water for 3h at room temperature, after which the plates are washed twice with sterile water and dried completely. To obtain single cells for microscopy the embryoid bodies are split into single cells using TripLE. After collecting the cell aggregates, they are washed in prewarmed PBS and treated for 3min with cold TripLE. Now the cells can be singularized via pipetting using Feeder's medium. Before plating, the cells have to be spun down and washed again with PBS. A density of 100.000 cells per cm² is optimal for microscopy purposes. After three days of culturing in monolayer medium with doxycycline,

the cells are fully differentiated. Depending on the number of cells, the medium must be refreshed every two to three days or if it turns yellow.

4.2.1.3. RNA transfection

RNA transfection experiments were performed using Lipofectamine messengerMAX, following the manufacturer's instructions. For filter experiments, 1 million cells were transfected with 5µg of A-tailed RNA reporter using 7.5µl Lipofectamine reagent in 250µl OptiMEM. One master mix was produced for multiple transfections and used for all reactions.

4.2.1.4. DNA transfection

DNA transfections were performed using JetPrime reagent following the manufacturer's instructions. One million cells were transfected in a 10cm cell culture dish with 10ml of growth medium using 10µg of plasmid DNA and 30µl of JetPrime reagent. One master mix was produced for multiple transfections and used for all reactions.

4.2.1.5. Generation of stable cell lines using the piggyback transposon system

Cells were transfected with a donor plasmid (piggyback donor) and a transposase plasmid (piggyback transposase) in a ratio of 5:1. After one day, the cells were treated with an appropriate antibiotic (200µg/ml hygromycin or 4µg/ml blasticidine) for five days or until the relative number of positive cells reached 100%. The percentage of positive cells was validated throughout the experiment using fluorescent reporter proteins (e.g. GFP) encoded on the donor plasmid.

4.2.1.6. Separation of subcellular compartments of N1E-115

For the starting culture, cells were growing 10cm cell culture dishes until they reach 50% to 70% confluency. After washing the cells with warm PBS, the medium was replaced by differentiation medium for one day. Meanwhile the 6-well cell culture inserts were coated from the outside using 10µg/ml laminin solution in PBS. For that, the filters are placed upside down on the lid of a 6-well dish and 500µl of coating solution is placed on top. After closing the 6-well-dish it was stored for 24h in the fridge. Before plating the cells, the coating solution was aspirated, the filters were washed in sterile water and placed in a 6-well cell culture dish with 2ml of differentiation medium. After one day of pre-differentiation the cells were carefully detached using PUCK's saline solution, washed once with warm PBS and placed inside the insert in another 2ml of differentiation medium. After 24h the cells built long protrusions through the filter, which attached to the outside.

For separating the compartments, the filters were placed in ice-cold PBS for 10s. Afterwards one of the two compartments was removed mechanically using a cotton swab dipped in ice-

cold PBS. Each filter has to be checked under the microscope to ensure efficient cleaning of the unwanted compartment. The remaining structures on the filter were extracted using either TriFast for RNA extraction or Tris buffered 8M Urea (protein extraction buffer). To gain efficient protein extraction, the filters were sonicated using a Bioruptor (15s ON, 45s OFF, high, 4 cycles) and the cell debris was subsequently sedimented via centrifugation (14.000g/3min/4°C), after which the supernatant was used for further analysis. The approximate yield per membrane is 15µg of RNA and 400µg of protein for the soma fraction and 400ng of RNA and 15µg of protein for the protrusion fraction.

4.2.1.7. Separation of subcellular compartments of Ascl1-iNs

One million cells were placed in 10ml of AK medium for pre-differentiation. After two days the formed embryoid bodies are washed with warm PBS and the cells are split into two 10cm dishes, each containing 10ml of AK medium supplemented with 3µg/ml doxycycline for two more days. Millicell cell culture inserts are coated for three hours at 37°C from the outside with Matrigel solution (1/30 in KnockOut DMEM medium). To do this, the inserts are placed upside down on the lid of a 6-well-dish, 500µl of Matrigel solution is placed on top of the filter and the 6-well-dish is closed again. Before plating the cells, the filters were washed with warm PBS to remove remaining coating solution and 2ml of Monolayer medium were placed below the filter into the 6-well dish. All embryoid bodies from one 10cm dish were placed into one 6-well-insert in 2ml of monolayer medium for differentiation. After three days of final differentiation the filters were processed to extract either soma or neurite fractions.

The differentiated cells on the membrane were cooled in ice-cold PBS prior to separation. The soma can now be torn from the neurites via pipetting vigorously with 1ml of ice-cold PBS. The soma is then collected via centrifugation (1000g/5min/4°C) and, meanwhile, the neurite sample is extracted. For this, the leftover soma fragments on the filter are cleaned using a cotton swab dipped in ice-cold PBS. To ensure a high purity of the sample, each filter has to be controlled using the light microscope for efficient removal of the soma. For RNA extraction the filters and the spun down soma fractions were put in TriFast. For protein extraction the filters and the spun down soma fractions were put in protein extraction buffer and sonicated using a Bioruptor (15s ON, 45s OFF, high, 4 cycles). The approximate yield of one filter is 2µg of RNA and 200µg of protein for each soma sample, and 200ng of RNA and 20µg of protein for each neurite sample.

4.2.2. Biochemical assays

4.2.2.1. Cloning

Cloning of the reporter plasmids was performed following the Gibson Assembly protocol. 10µg of the initial vector was restricted in a volume of 50µl, using 5U of two non-compatible restriction enzymes from NEB at a temperature of 37°C overnight. 2.5U of alkaline phosphatase was added for two more hours. The restricted vector was purified from a 1% agarose gel using GeneJet gel extraction kit (K0692) and the construct was eluted in 30µl. The inserts were PCR amplified using primers with compatible overhangs and also purified agarose gel. The final reaction was assembled in a 1:5 molar ratio of vector to insert and mixed 1:1 with homemade Gibson assembly mix. Gibson assembly reactions were performed at 50°C for 100min. For each reaction a negative control containing no insert was run. The assembled vector was transformed into chemical competent Mach1 *E. coli* strains or, if very high transformation rates were needed, it was dialyzed against ultrapure water and transformed into electro competent Endura cells (60242-1). Homemade chemical competent cells were thawed on ice for 20min prior to use. 2µl of the Gibson assembly mix was mixed with 50µl of bacterial suspension and incubated on ice for 10min. The bacterial suspension was then heat shocked at 42°C for 60s and cooled down on ice again immediately. Cultures were recovered in 1ml LB medium for 45min prior to striking out on solid LB plates supplemented with the appropriate antibiotic. Electro transformations were performed using the Gene Pulser II electroporator (Bio-Rad 165-2105) and Lucigen Endura electrocompetent cells following manufacturers instructions. 1mm cuvettes were used with a pulse of 10µF, 600Ohms and 1,800V. The efficiency of each clone was calculated by comparing the positive reaction with the negative control reaction lacking the insert. Positive clones were either grown in 5ml of LB medium supplemented with an appropriate antibiotic for a test culture or immediately mixed after recovery and grown in 500ml of LB medium with appropriate antibiotic for a plasmid library.

<u>Isothermal Start Mix</u>

1.5 g PEG8000

3 ml 1 M Tris--HCl, pH 8.0

150 µl 2 M MgCl2

Total: 3150 µl

<u>2X Gibson Assembly Master Mix</u>

405 µl Isothermal Start Mix (RT)

25 µl 1 M DTT (4°C)

50 µl 10 mM dNTPs (-20°C)

50 µl NAD+ (NEB # B9007S) (-80°C)

1 µl T5 exonuclease (NEB # M0363S – 10 U/µl) (-20°C)

31.25 µl Phusion High Fidelity DNA Polymerase (NEB # M0530S – 2 U/µl) (-20°C)

250 µl Taq Ligase (NEB Cat. # M0208S – 40 U/µl) (-20°C)

437.75 µl H2O

Total: 1250 µl

4.2.2.2. Preparation of samples for liquid chromatography tandem mass spectrometry (LS MS/MS)

All samples for LS MS/MS were prepared in protein extraction buffer (0.1M Tris pH:8, 8M Urea). Total cells were disrupted using the Bioruptor with the settings 15s ON, 45s OFF, high, 4 cycles. Afterwards the cell debris was spun down for 10min at 14,000g and 4°C. The supernatant was then transferred to a fresh tube and stored at -80°C until digestion.

4.2.2.3. Stable isotope labelling by amino acids (SILAC)

To gain more quantitative results from mass spectrometry it is necessary to label the proteins beforehand. Here we labelled cells with amino acids, which differ from each other in weight due to different incorporated isotopes. We used L-lysine monohydrochloride and L-arginine monohydrochloride incorporation for light labelling and "Lys 8" (L-lysine-$^{13}C_6$ $^{15}N_2$ monohydrochloride) and "Arg 10" (L-arginine-$^{13}C_6$ $^{15}N_4$ monohydrochloride) incorporation for heavy labelling. Cell culture medium deprived from natural lysine and arginine is supplemented with these amino acids and used for growing the cultures. Cells were grown for at least five passages in labelled medium to avoid all residues of natural amino acids, which would have interfered with the quantification. Prior to the experiment the labeling efficiency was tested via MS. For differentiation testing, two different cultures of cells were prepared and differentiated in either heavy or light labelled medium. After protein purification via sonication (as described above) and concentration measurement via Qubit protein assay, equal amounts were mixed and tested in pairs, as heavy labelled non-differentiated protein extracts mixed with light labelled differentiated protein extracts and vice versa. The label swap experiment is crucial to exclude a bias of the labelled media on the results and, for further analysis, only proteins which showed a significant enrichment in both experiments were taken. For localization experiments the labelled cells were differentiated in SILAC labelled monolayer medium and compartment separation was performed. Protein samples from light labelled soma fractions were pooled with the same amounts of proteins coming from heavy labeled neurite

fractions and vice versa. All protein extracts were prepared using protein extraction buffer and digested and analyzed by MS/MS.

4.2.2.4. Pulsed stable isotope labelling by amino acids (pSILAC)

For pSILAC experiments, non-labelled cells (light) were cultured for a certain time in either heavy or medium labelled medium. The incorporation of heavy and medium labelled amino acids is taken as a measure for protein production in the labelling timeframe. Medium pSILAC medium was supplemented with "Lys 4" (4,4,5,5,-D_4-L-lysine monohydrochloride) and "Arg 6" (L-arginine-$^{13}C_6$ monohydrochloride), while heavy labelled pSILAC medium was supplemented with "Lys 8" (L-lysine-$^{13}C_6$ $^{15}N_2$ monohydrochloride) and "Arg 10" (L-arginine-$^{13}C_6$ $^{15}N_4$ monohydrochloride). The pulse length was set for two hours. Prior to pulse labelling, the cells were washed twice with warm PBS to remove any residual non-labelled medium. Protein samples from separated subcellular compartments were prepared in protein extraction buffer via sonication, mixed in equal amounts and submitted to digestion and MS.

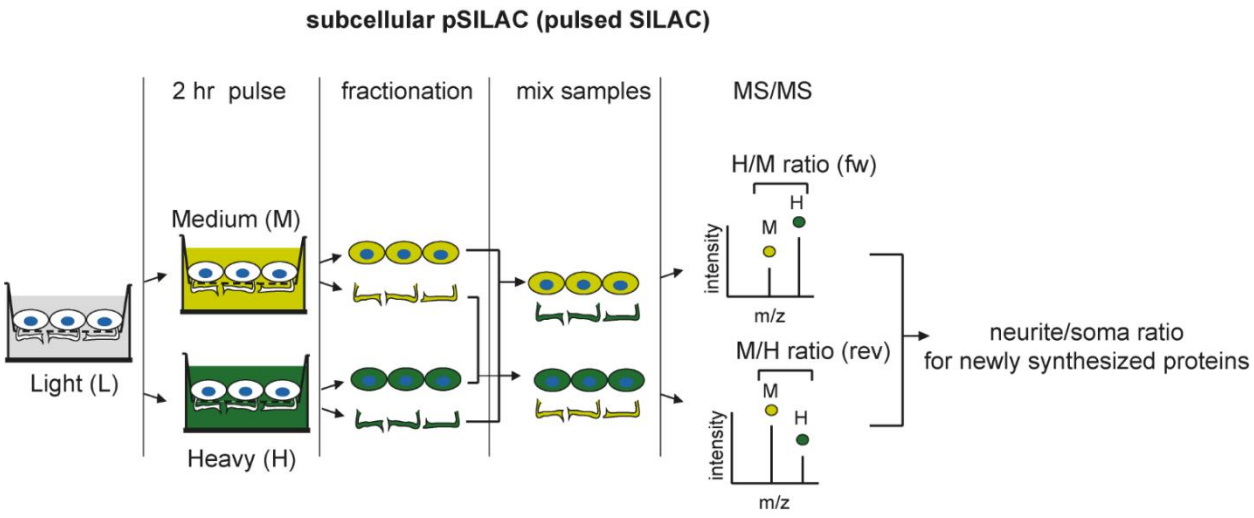


Fig. 4 pSILAC. Schematic presentation of the pSILAC experiments.

4.2.2.5. Quantitative bio-orthogonal noncanonical amino acid tagging with stable-isotope labelling by amino acids (QuaNCAT)

In QuaNCAT experiments the pulse labelling with amino acids is combined with bio-orthogonal amino acid tagging and L-azidohomoalanine. This nontoxic alanine orthologue can be used to create covalent bindings via Click-iT ® chemistry. This click chemistry is used in the assay to enrich newly synthesized proteins via bead purification and therefore reduces the pulse length dramatically. Medium QuaNCAT medium was supplemented with "Lys 4" (4,4,5,5,-D_4-L-lysine monohydrochloride),"Arg 6" (L-arginine-$^{13}C_6$ monohydrochloride) and AHA (L-azidohomoalanine), while heavy labelled pSILAC medium was supplemented with "Lys 8" (L-lysine-$^{13}C_6$ $^{15}N_2$ monohydrochloride), "Arg 10" (L-arginine-$^{13}C_6$ $^{15}N_4$ monohydrochloride) and AHA (L-azidohomoalanine). The pulse length of this experiment was set for 30min. TheClick-iT protein purification kit (Thermo Fisher Scientific C10416) was used for Click-iT protein purification following the manufacturer's protocol. Since the amount of proteins coming from neurite samples is always limiting, 21 filters were pooled per neurite sample. To extract those in the required 400µl of lysis buffer, batches of 7 membranes were sonicated and the supernatant buffer containing the proteins was transferred to a new tube containing the next 7 filters and so on. After preparing protein samples the protein extracts were mixed again in complementary pairs (heavy soma with light neurites and vice versa) and Click-iT enrichment for newly synthesized proteins was performed following the manufacturer's protocol. Afterwards protein pools were submitted to digestion and MS analysis.

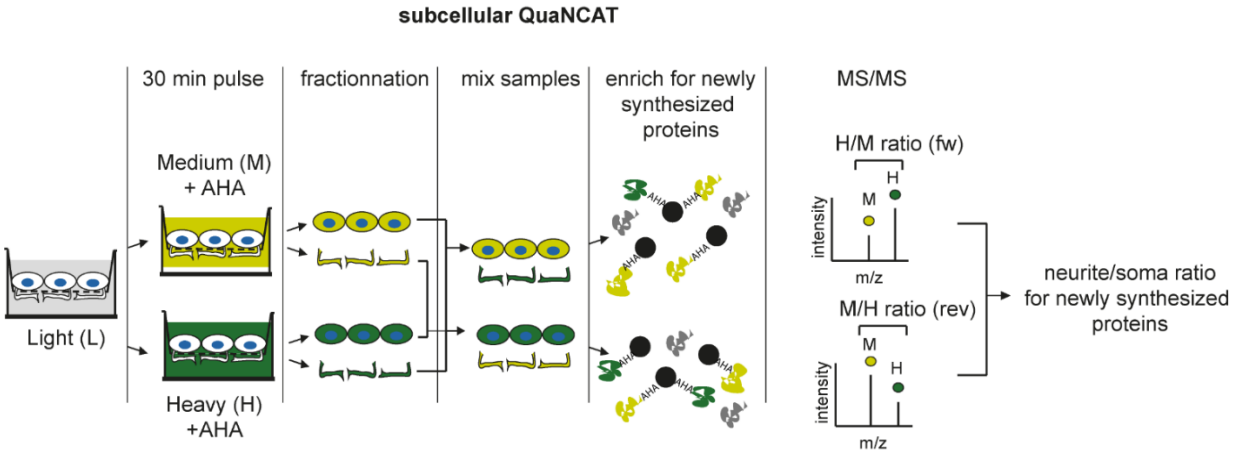


Fig. 5 QuaNCAT. Schematic presentation of the QuaNCAT experiments.

4.2.2.6. SDS-PAGE

SDS-PAGE was performed using the Bio-Rad Trans-Blot Turbo device (Bio-Rad 1704155). At first the separating gel was assembled and poured into the gel-holding device. To ensure a linear border to the stacking gel and no air bubble inclusions it was topped up with water until it solidified (30min). Afterwards the stacking gel was poured on top and the device was closed with the comb for another 30min. Once ready, the gel was mounted into the running device and SDS running buffer was filled into the outer and inner part of the assembly. Protein samples were mixed 1:5 with Laemmli loading buffer and heated for 5minto 95°C to ensure protein denaturation. Before loading the gel, all wells were flushed with a syringe. 25µl of protein solution was loaded into each well and the gel was run at 80V until the bromophenol blue border reached the separation gel and then at 120V until the bromophenol blue border exited the gel on the bottom. After disassembling the running device, the gel was either used for a Coomassie Brilliant Blue staining or subsequently for western blotting.

SDS loading buffer 5X (aliquots in freezer 5)

component	final concentration	units	concentration in stock	units	units	added amount	units
Tris-HCl (pH 6.8)	300.0	mM	1500	mM	ml	8.0	ml
glycerol	50.0	%	100	%	ml	20.0	ml
SDS	10.0	%	100	%	g	4.0	g
14.3 M β-2-mercapto-ethanol	5.0	%	100	%	ml	2.0	ml
bromphenol blue	0.1	%			mg	40.0	mg
H_2O						add to 40.0 ml	
final volume					**ml**	**40.0**	ml

SDS-PAGE running buffer (Tris-Glycine/SDS buffer pH 8.3) – shared stock

	1X		10X				
component	final concentration	units	final concentration	units	added amount for 10X	units	M (molar mass), g/mol
Tris-base/Trizma	0.025	M	0.13	M	30.3	g	121.14
Glycine	0.190	M	0.95	M	144.1	g	75.06
SDS	0.1	%	1.00	%	10.0	g	
H_2O					0.95	L	
final volume					**1.00**	**L**	

	1x mini gel		2x mini gels		3x mini gels		4x mini gels	
7.5% separating gel:	6	ml	12	ml	18	ml	24	ml
H2O	2.9	ml	5.8	ml	8.7	ml	11.6	ml
Acrylamide:Bis 30% 37.5:1	1.5	ml	3	ml	4.5	ml	6	ml

1.5 M Tris pH 8.8	1.5	ml	3	ml	4.5	ml	6	ml
20% SDS	0.03	ml	0.06	ml	0.09	ml	0.12	ml
10% APS	0.06	ml	0.12	ml	0.18	ml	0.24	ml
TEMED	0.005	ml	0.010	ml	0.015	ml	0.020	ml

10% separating gel:	6	ml	12	ml	18	ml	24	ml
H2O	2.4	ml	4.8	ml	7.2	ml	9.6	ml
Acrylamide:Bis 30% 37.5:1	2	ml	4	ml	6	ml	8	ml
1.5 M Tris pH 8.8	1.5	ml	3	ml	4.5	ml	6	ml
20% SDS	0.03	ml	0.06	ml	0.09	ml	0.12	ml
10% APS	0.06	ml	0.12	ml	0.18	ml	0.24	ml
TEMED	0.005	ml	0.010	ml	0.015	ml	0.020	ml

12.5% separating gel:	6	ml	12	ml	18	ml	24	ml
H2O	1.9	ml	3.8	ml	5.7	ml	7.6	ml
Acrylamide:Bis 30% 37.5:1	2.5	ml	5	ml	7.5	ml	10	ml
1.5 M Tris pH 8.8	1.5	ml	3	ml	4.5	ml	6	ml
20% SDS	0.03	ml	0.06	ml	0.09	ml	0.12	ml
10% APS	0.06	ml	0.12	ml	0.18	ml	0.24	ml
TEMED	0.005	ml	0.010	ml	0.015	ml	0.020	ml

15% separating gel:	6	ml	12	ml	18	ml	24	ml
H2O	1.4	ml	2.8	ml	4.2	ml	5.6	ml
Acrylamide:Bis 30% 37.5:1	3	ml	6	ml	9	ml	12	ml
1.5 M Tris pH 8.8	1.5	ml	3	ml	4.5	ml	6	ml
20% SDS	0.03	ml	0.06	ml	0.09	ml	0.12	ml
10% APS	0.06	ml	0.12	ml	0.18	ml	0.24	ml
TEMED	0.005	ml	0.010	ml	0.015	ml	0.020	ml

5% stacking gel:	3	ml	6	ml	9	ml	12	ml
H2O	2.2	ml	4.4	ml	6.6	ml	8.8	ml
Acrylamide:Bis 30% 37.5:1	0.5	ml	1	ml	1.5	ml	2	ml

1.5 M Tris pH 6.8	0.25	ml	0.51	ml	0.76	ml	1.02	ml
20% SDS	0.015	ml	0.03	ml	0.045	ml	0.06	ml
10% APS	0.03	ml	0.06	ml	0.09	ml	0.12	ml
TEMED	0.003	ml	0.006	ml	0.009	ml	0.012	ml

4.2.2.7. Western blotting

All western blots were performed using the semi dry transfer method. SDS-PAGE were assembled and run on the Bio-Rad Transblot Turbo station. The Gel was equilibrated in transfer buffer for 10min. The PVDF membrane was activated in pure methanol for 15s and also equilibrated in transfer buffer for 10min. For each blot two stacks of 7 Bio-Rad filter pads each were soaked in transfer buffer as buffer reservoirs. The transfer sandwich was assembled by layering onto the bottom of the transfer cassette (anode) one stack of buffer-soaked filter pads, the PVDF membrane, the equilibrated gel and, on top, another stack of buffer-soaked filter pads. All remaining air bubbles in this stack were removed and the transfer cassette was closed. The transfer was performed using 25V and 1A for 30min. The efficiency of the transfer was monitored via a prestained protein marker (Thermo Fisher Scientific 26619) and also via Coomassie staining the remaining gel. Prior to immunostaining, the membrane was blocked in blocking buffer (3% skim milk powder solution in TBS) for 30min of shaking. Primary antibody incubation was performed overnight at 4°C in blocking buffer with an appropriate antibody concentration. After the primary incubation, the membrane was washed three times with TBS-T for 15min each. Then the secondary antibody, coupled to horseradish peroxidase, was incubated on the membrane for one hour, shaking at room temperature. The secondary antibody was washed away with TBS-T three times, for 15min each wash, and pictures were made on the ImageQuant LAS4000 using homemade enhanced chemiluminescence (ECL) reagents.

Combine:

- 10 ml Tris pH 8.5
- 2,6 µl H_2O_2
- 25 µl coumaric acid
- 50 µl luminol

4.2.2.8. RNA extraction

RNA extractions were performed with TriFast following the manufacturer's protocol. All concentrations were measured using the Qubit RNA high sensitivity kit following the

manufacturer's protocol. All quality control measurements were performed using either the Bioanalyzer RNA nano or RNA pico kit, following the manufacturer's protocol and spectrophotometer measurements for phenol, protein or salt contaminations. RNA was always stored in appropriate aliquots in DEPC treated sterile water at -80°C.

4.2.2.9. cDNA synthesis and quantitative real time PCR (qRT-PCR)

Prior to cDNA synthesis, high quality RNA was treated with DNaseI following the manufacturer's protocol. cDNA synthesis was performed using the Maxima cDNA synthesis kit following the manufacturer's protocol. All qRT-PCR reactions were performed using the Bioline SensiFast Sybr NoRox kit following the manufacturer's protocol. Reactions were assembled in 20µl volume and technical triplicates as well as biological duplicates. For each primer pair a no template control was run in triplicates and for each primer-cDNA combination a no reverse transcription control was performed to validate the DNase treatment. Each reaction was performed using 2ng of diluted cDNA. Reactions were run on a BioRad CFX 96 touch qRT-PCR device and results were analyzed using the Bio-Rad CFX Manager version 3.1. For new primer pairs the amplicons were validated with a melting gradient and via agarose gel electrophoresis once, and the amplification efficiency was validated via cDNA dilution gradient. Fold change differences were calculated using the $\Delta\Delta C_T$ method using appropriate control genes. All error bars in qRT-PCR plots represent the standard deviation of biological replicates.

4.2.2.10. NanoString

NanoString experiments were conducted following the manufacturer's protocol. All Experiments were performed in duplicates using 500ng total RNA as input material. Results were analyzed using the nSolver software using internal controls for normalization.

4.2.2.11. Total RiboZero RNA sequencing

Total RiboZero RNA libraries were produced following the Illumina TruSeq stranded total RNA library protocol. Prior to the library production, all RNA samples were quality controlled for degradation with the Bioanalyzer and for contamination with the spectrophotometer. Concentrations throughout and prior to the protocol were measured via Qubit RNA high sensitivity assay. ERCC RNA mixes were spiked into all total RiboZero libraries. Ribosomal depletion was tested immediately after it was performed via Bioanalyzer and prior to sequencing using qRT-PCR. The amplification cycle number in the library generation was chosen by diagnostic qRT-PCR with a 1/10 fraction of the sample and set for the $C_T + 2$. Libraries were run on a NextSeq500 or HiSeq4000 following the manufacturer's protocol. All

libraries were quality controlled for size and concentration via Qubit DNA high sensitivity kit and Bioanalyzer DNA1000 chip.

4.2.2.12. Small RNA sequencing

Small RNA libraries were produced following the Illumina TruSeq Small RNA Library Prep Kit guide. Prior to the library production, all RNA samples were quality controlled for degradation with the Bioanalyzer and for contamination with the spectrophotometer. Concentrations throughout and prior to the protocol were measured via Qubit RNA high sensitivity assay. Small RNA libraries were all purified from TBE gels and all libraries were quality controlled for size and concentration via Qubit DNA high sensitivity kit and Bioanalyzer DNA1000 chip .

4.2.2.13. In vitro transcription and A-tailing

Fragments coming from a mixture of RNA-seq libraries were PCR amplified for 10 cycles using oligos, which add a T7 RNA polymerase promotor sequence. *In vitro* transcription was performed with mMESSAGE mMACHINE T7 in vitro transcription kit following the manufacturer's protocol. RNA was purified using RNeasy column purification following the manufacturer's protocol and A-tailed with Poly(A) Tailing Kit following the manufacturer's protocol. Prior to usage the RNA was column purified again using RNeasy columns.

4.2.2.14. Modified RNA element selection assay (RESA)

The RNA element selection assay (RESA), published by Yartseva et al. [284], is a method to identify regulatory RNA elements out of a fragmented reporter library. It is based on the introduction of a pool of fragmented RNA elements into a test system and analyzing the re-purified reporter library via next generation sequencing after a period of time. Changes in abundancy of specific sequences can be used to identify regulatory sequences that stabilize or destabilize the reporter in this system and timeframe. In order to use the theory in our test system several things were changed in the experimental setup; the reporter library was incorporated stably into the genome of our cell line and instead of a temporal effect on stabilization, the effect on localization of the transcript was analyzed.

384 transcripts that show significant RNA or protein enrichment in our dataset were chosen as reporter candidates. Additionally, 48 transcripts that show enrichment in Cajigas et al. [246] were chosen as control elements. Primers to amplify the 3' UTRs of these transcripts were designed using primer3 software (see material and methods - oligos). PCR amplification of 3' regions was performed with Phusion high fidelity polymerase in a 20µl volume and 96-well format using a temperature gradient from the highest to the lowest Tm of the primers on the

plate. 5µl off each reaction was used to validate the product by size via electrophoresis on a 1% agarose gel. PCR reactions that did not work were one repeated using the GC-buffer. PCR reactions that gave multiple amplicons were run on a gel and the correct band was excised and gel purified. The amount of amplicon was estimated based on the thickness of the band and either 5µl, 10µl or 15µl of the reactions were pooled to create the reporter library. The pool of fragments was batch purified using AmpureXP beads and eluted in water. 10µg of the reporter pool was used for random fragmentation via sonication using a Bioruptor with the settings 30s ON, 30s OFF, high, 40 cycles. After fragmentation the fragments were blunt ended using the Klenow fragment, adding 6µl of T4 ligase buffer, 1µl of dNTPs and 3µl of Klenow fragment for one hour at 37°C. The fragmented and blunt ended material was run on a Mini-PROTEAN gel and fragments with a size between 100 and 150nt were excised from the gel.

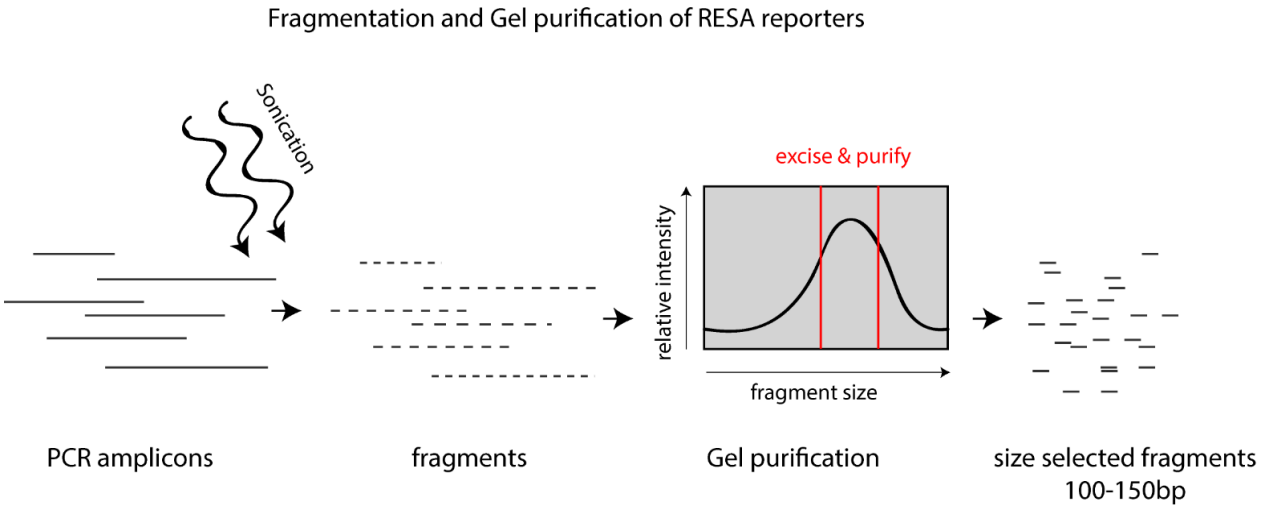


Fig. 6 Fragmentation and gel purification of RESA fragments

The gel slice was crushed and diluted with 50µl of DEPC treated water and rotated overnight at 4°C for elution. Afterwards the gel residues were spun down and the dsDNA fragments from the supernatant were purified using AmpureXP beads and eluted in 17.5µl of water. 2.5µl of resuspension buffer and 12.5µl of A-tailing mix, both from the Illumina TruSeq stranded total RNA-seq kit, were added and the tubes were put on the thermal cycler running the A-tailing program from the kit (preheat lid to 100°C, 37°C/30min, 70°C/5min, 4°C/hold). Following the adapter ligation procedure from the Illumina TruSeq stranded total RNA-seq protocol 2.5µl of resuspension buffer, 2.5µl of ligation mix and 2.5µl of barcoded adapter were added and the whole sample was mixed intensively via pipetting and incubated on a thermal cycler for 10min at 30°C. 5µl of stop ligation buffer was added and the adapter ligated fragments were purified twice with AmpureXP beads in a ratio of 1:1 (sample : beads). The last elution step was performed in 32.5µl of resuspension buffer. A diagnostic qRT-PCR was run to estimate the

amount of adapter ligated library and the final PCR amplification was performed using the Phusion high fidelity polymerase with HF buffer in a 50µlvolume and primers 2425 and 2426 with C_t+2 cycles to add overhangs for Gibson assembly into the AgeI restricted piggyBAC transfer vector. All steps were validated using the Bioanalyzer DNA1000 chip.

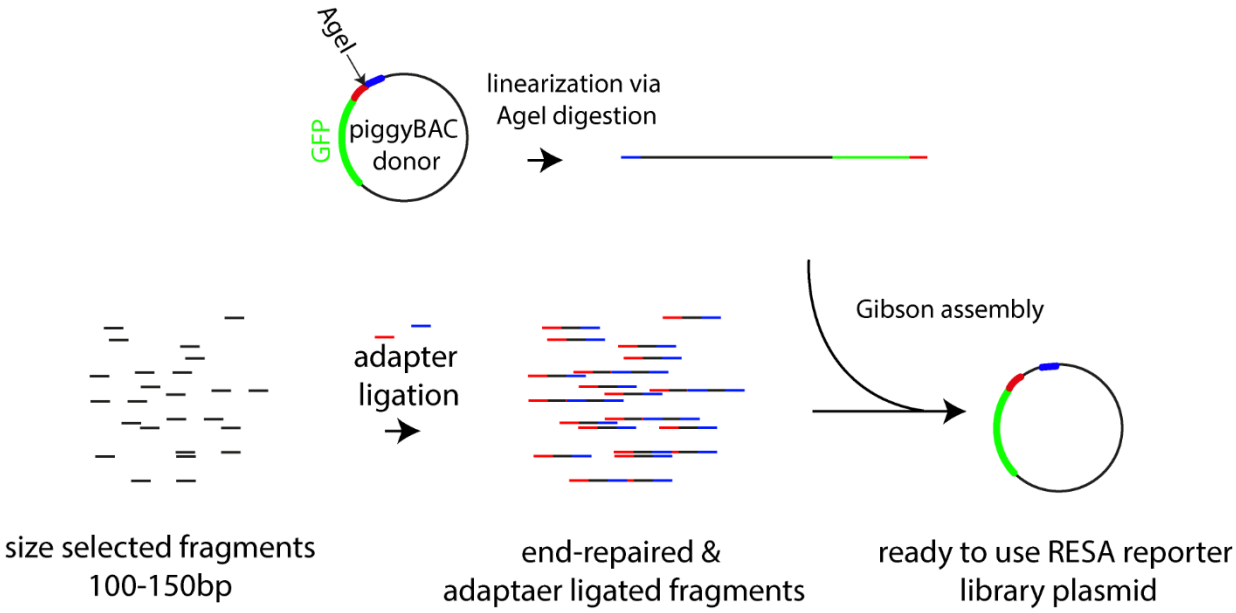


Fig. 7 Cloning of RESA fragments

The vector backbone for cloning was prepared via AgeI digestion of 10µg of piggyBAC transfer vector overnight at 37°C. The linearized product was treated with FastAP for three hours at 37°C to prevent re-ligation and purified via gel electrophoresis on a 1% agarose gel and eluted in 30µl of water. 10µl of this linearized vector was mixed with either 15µl of insert or water for the negative control and 25µl of 2x homemade Gibson mix and incubated for 2h at 50°C. The complete Gibson assembly reaction was dialyzed against water and transformed in 10 reactions into electro competent Endura cells (Lucigen 60242-1). After recovery of the cells for a maximum of 30min in supplemented recovery medium, a serial dilution of the mix starting from 10µl of undiluted bacterial suspension was plated on LB ampicillin plates to estimate the number of c.f.u. in the experiment. The number of c.f.u. is taken as a proxy for the number of single ligation events and therefore can be used to calculate the estimated complexity of the final library. The rest of the bacterial suspension is used to inoculate 500ml of LB supplemented with ampicillin. This culture is used to prepare the plasmid library via plasmid DNA isolation using the Nucleobond Xtra Midi Plus kit (Macherey Nagel MN/00740412/000050).

To prepare a reporter cell line, 160µg of the piggyBAC transfer vector carrying the library was mixed with 40µg of piggyBAC transposon vector and transfected into 20 million Ascl1-mESC using JetPrime transfection, following the protocol outlined above. Cells were treated with 200µg/ml of hygromycin for three days until the fraction of GFP positive cells was >90%. After one day of recovery from antibiotic selection the cells were differentiated using the protocol outlined above and distributed onto 54 filters coated with Matrigel. Filter separation was performed as stated above and the neurites and soma of 18 filters are always pooled as one replicate. RNA was purified as stated previously and was tested for integrity, purity and quantity via Bioanalyzer, spectrophotometer and Qubit RNA high sensitivity assay. 50ng of each RNA sample was taken as input for cDNA synthesis using the Maxima cDNA synthesis kit and qRT-PCR analysis using primers for the library, rRNA and Tagln to estimate the amount of library and the efficiency of separation. 1.8ug of extracted RNA is RQ1-DNase treated and, after AmpureXP bead purification, suspended in 13.1µl of water was taken as input for targeted cDNA synthesis using the TGI-RT. 1µl of 10µM primer 2317 solution was added and the mixture was heated to 70°C for 10min and cooled on ice immediately. 4µl of TGI-RT buffer was added together with 80U of the TGI-RT enzyme. The reaction was left for 30min at room temperature for priming. 2.5µl of dNTPs was added and the reverse transcription reaction was performed for two hours at 60°C. To terminate the reaction and separate the TGI-RT from the constructs it was necessary to treat the mixture with 1µl of 5N NaOH for 3min at 95°C, after which it was neutralized with 1µl of 5N HCl on ice. cDNA was bead purified using the AmpureXP beads and eluted in 30µl of water. 2µl of this mixture was taken as input for another diagnostic qRT-PCR to estimate the correct cycle number for the final amplification. The final amplification of the library for sequencing was also used to introduce different barcodes to the different samples using primer pairs 1453/2317, 1454/2317, 1455/2317, 1456/2317, 1457/2317 and 2317/2318 for samples Neurites1, Neurites2, Neurites3, Soma1, Soma2 and Soma3, respectively. PCR reaction was performed in 50uL volume with Phusion high fidelity polymerase in HF buffer and diagnostic PCR C_T+2 as cycle number. Once ready, the libraries were purified with AmpureXP beads and mixed in equimolar amounts for sequencing.

4.2.3. Microscopy

4.2.3.1. Immunohistochemistry

For immunostainings of single neurons, cells were differentiated on coated cover slides. The coating for N1E-115 cells was 1ug/ml laminin solution in PBS for three hours at room temperature. The laminin solution was washed away shortly before seeding the cells. The

coating of glass cover slides for Ascl1-mESC was performed for 3h at room temperature with a 0.1mg/ml Poly-L-Ornithine solution in water. The coating solution was washed away with water and the glass slides were dried. N1E-115 cells were seeded undifferentiated on the laminin coated glass slides and differentiated attached to the slide. Ascl1-mESC embryoid bodies were prepared as described above. After two days of culturing in AK supplemented with 3µg/ml doxycycline, embryoid bodies from one 10cm dish were washed gently with warm PBS and incubated in 1ml TripLE for 3min at room temperature. Cells were separated via pipetting using 4ml of Feeders medium. The cell suspension was spun down at 1,500g for 5min and the cells were suspended in monolayer medium. Cell concentration was counted using the Luna cell counter (Logos Biosystems CTM0011) and plated in a density of 150,000 cells per cm^2 in monolayer medium. The medium was changed once on the third day of differentiation on the slide. Cells were processed on the fourth day of culturing them in monolayer. N1E-115 cells were processed on the third day of differentiation on the cover slide. The staining was performed as followed:

- Cells washed gently three times with cold PBS
- Fixation for 10min at room temperature in 4%PFA in PBS
- Cells washed gently three times with cold PBS for 5min each
- Permeabilization at room temperature with PBS supplemented with 0.2% Triton X-100
- At this step cells were stored in the fridge for up to one month in PBS or 70% EtOH
- Blocking with PBS supplemented with 0.05% TWEEN and 20% Roche reagent
- First antibody incubation at 4°C overnight in blocking solution with recommended antibody concentration
- Cells washed three times with PBS for 5min each
- Second antibody incubation at room temperature in the dark in blocking solution with recommended concentration
- Cells washed three times with PBS supplemented with 0.05% TWEEN for 5min each
- Mount cells with ProLong Gold on a glass slide holder
- Dry samples for one day at room temperature in the dark

Cells grown on the filter were treated as described above until the mounting step. One drop of mounting medium was placed on the glass slide holder and the filter was excised from the plastic holder and placed on the mounting medium. Another drop of mounting medium was placed on top of the filter and the coverslip was placed on top of the sandwich. Filter samples were also dried for one day at room temperature in the dark. Pictures were taken using a Laica SP8 confocal microscope.

5. Results

5.1. Characterization and differentiation of the cell line N1E-115

N1E-115 neuroblastoma cells were differentiated by serum starvation following the protocol from Pertz et al [282]. The neuronal phenotype was validated via immunofluorescence (Fig. 8a). These cells were used as neuronal model systems in diverse studies [285]–[289]. To confirm the neuronal phenotype on the transcriptome level a RiboZero total RNA sequencing experiment was performed with RNA extracted on d0 and d3 of the differentiation. Using DeSeq2, 20,606 genes were identified in total with more than two reads per condition, of which 1,175 transcripts showed an upregulation of twofold or more with a p-val below 0.05, while 219 transcripts were downregulated twofold or more with a p-val below 0.05 (Fig 8b). Gene ontology (GO) term enrichment of the upregulated genes (differentiated N1E-115/undifferentiated N1E-115 > 4) was analyzed using the Panther software [290] **(Figure 1)**. A gene ontology overrepresentation test for biological processes showed a clear enrichment for neuronal associated terms, such as: regulation of axonogenesis, neurogenesis, synaptic vesicle formation and synaptic plasticity as well as several terms related to intracellular protein and organelle localization (Fig 8c).

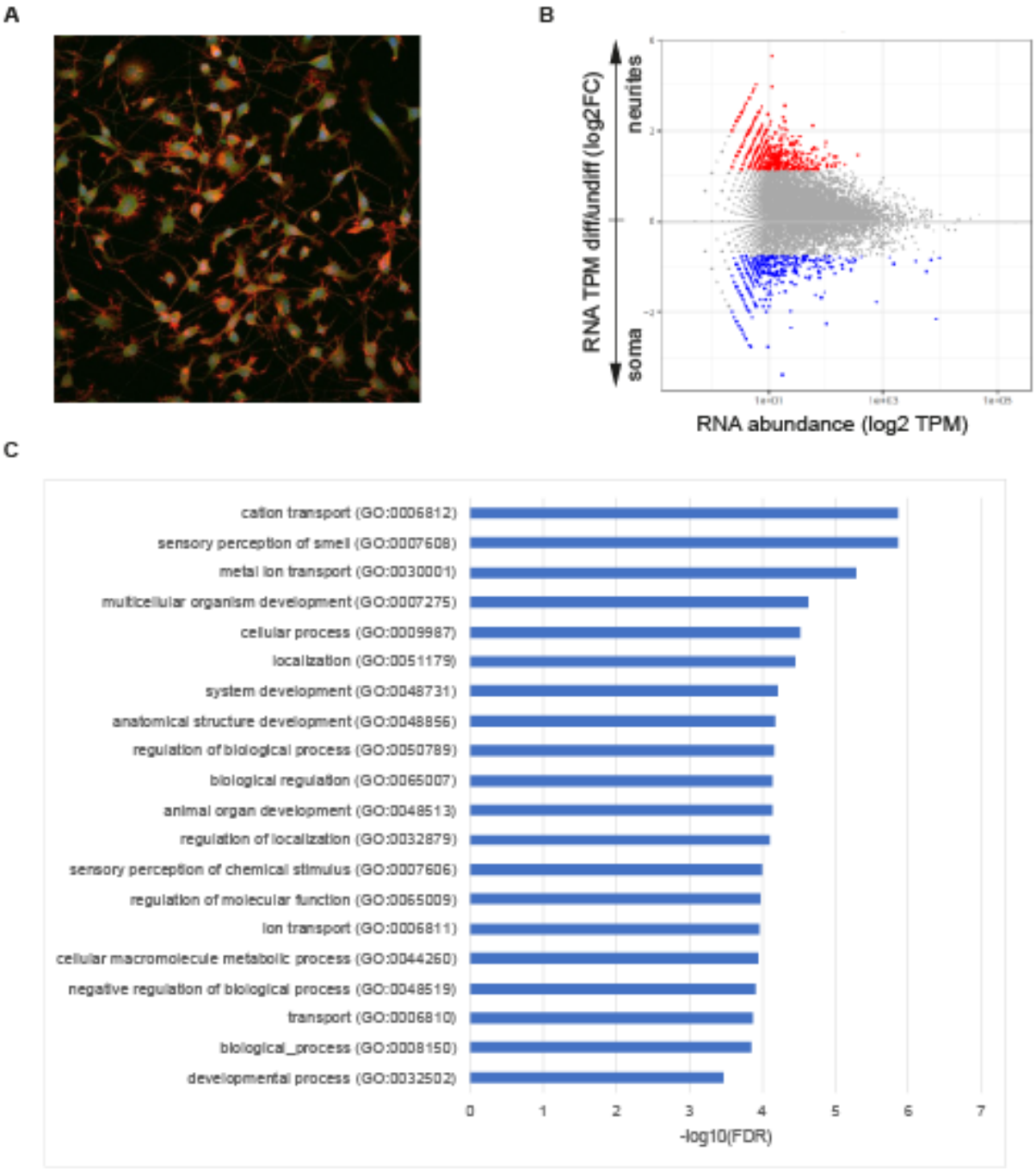


Fig. 8 Characterization of the Neuroblastoma cell line N1E-115

A Immunofluorescence of two days differentiated N1E-115 cells. Differentiation was induced via serum starvation. DAPI (blue), Actin (red) and Tubulin (green). **B** Batch RiboZero RNA-seq of undifferentiated and differentiated cells. RNA samples were prepared after two days of treatment. Primary analysis of the data was performed by Vedran Franke. **C** Statistical enrichment of gene ontologies during differentiation reveals a statistical overrepresentation of terms related to neuronal development. Shown are the top 20 results of the statistical overrepresentation test.

5.2. Characterization and differentiation of Ascl-mESC into Ascl1-iNs

Ascl1-iNs were used as a second test system throughout this study. These cells were differentiated into neurons from mESCs, expressing the neurogenic transcription factor Ascl1 under the doxycycline-inducible promotor (Ascl1-mESC). Ascl1 is expressed endogenously early during neuronal differentiation. It is a pioneer proneural transcription factor that drives differentiation towards induced neurons (Ascl1-iNs) that are post-mitotic and express neuronal markers.

It has been shown that Ascl1 alone is sufficient to generate functional neurons from mouse and human ESCs and astroglia [291], [292]. Homogeneity of the neuronal cell population is the critical precondition for performing neuronal localization experiments. Since the Ascl1-iN cells showed high homogeneity they were chosen for further experimental manipulation.

I confirmed the neuronal identity of these cells by SILAC (stable isotope labeling by amino acids in cell culture), comparing uninduced Ascl1-mESC and Ascl1-iNeurons (Fig. 9a). In these experiments differences in protein abundance between samples are detected using isotopic non-radioactive labelling [293]. As expected, GO term overrepresentation analysis [290] showed that genes upregulated upon differentiation (iNeurons/mESC > 4) are associated with neuronal functions. Additionally, immunofluorescence using markers for neurofilament and tubulin showed a clear neuronal phenotype. (Fig. 9c).

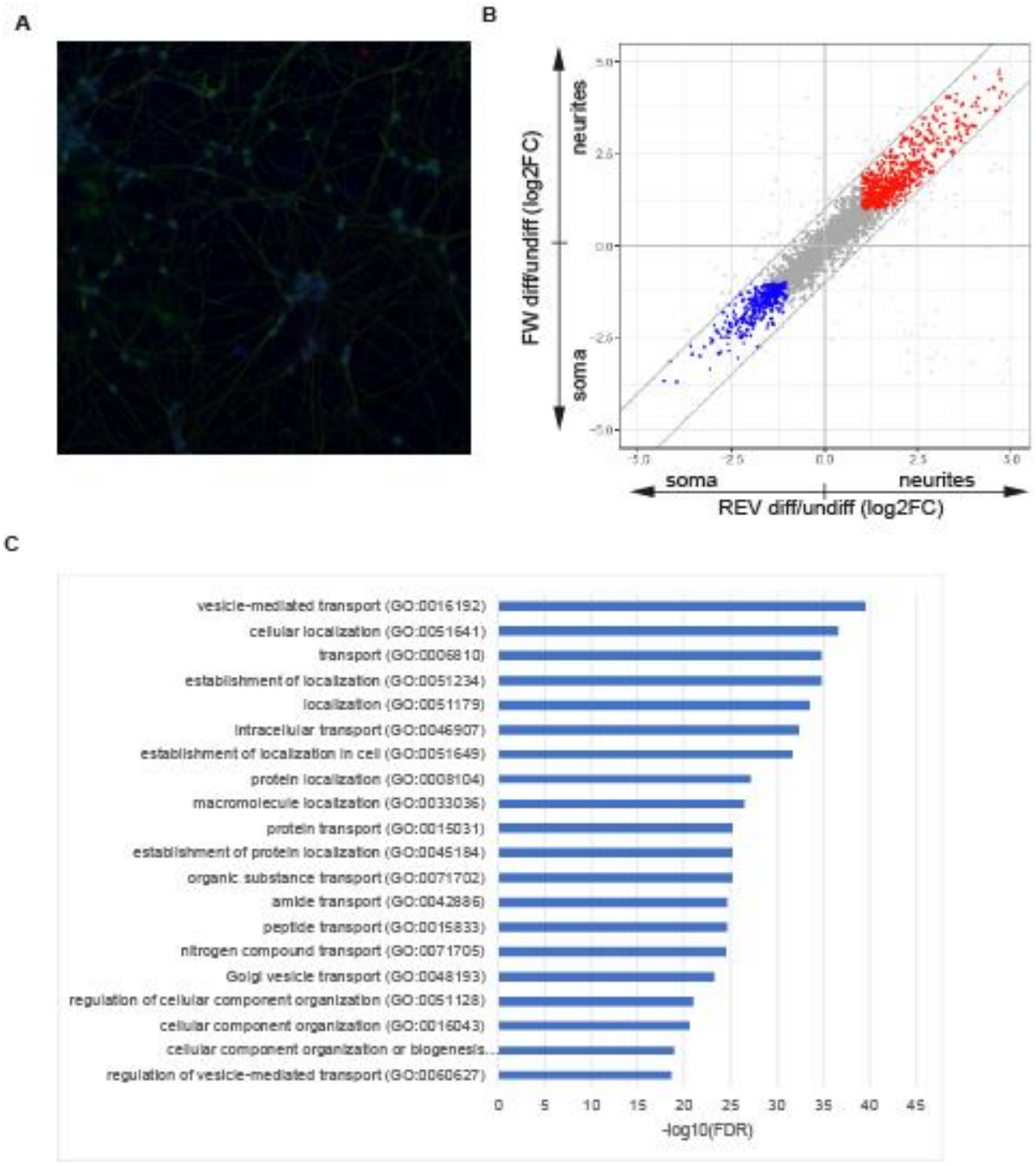


Fig. 9 Characterization of Ascl1-iNs

A Immunofluorescence pictured five days after Ascl1 induction. Neurofilament (green), Tubulin (red), DAPI (blue). **B** Stable Isotope Labelling by amino acids followed by MS/MS. Upregulated proteins during differentiation are highlighted in red, downregulated proteins are highlighted in blue. Primary analysis of the data was performed by Koshi Imami. **C** Upregulated proteins are associated with neuron related gene ontologies.

5.3. The local transcriptome and proteome of neuronal cells revealed by spatial omics

To identify proteins and mRNAs that are asymmetrically localized between the neurites and soma of neuronal cells, a modified Boyden chamber assay that allows the separation of distinct cellular compartments was employed.

To isolate neurite outgrowths and soma for biochemical applications, ASCL1-iNeurons and N1E-115 neuroblastoma cells were differentiated and maintained on a porous membrane in such a way that cell bodies stayed on the upper side of the membrane. The coating agent on the lower side of the membrane provided cues for neurites to grow through the pores and attach onto the other side of the membrane [282], [294].

Indeed, neurofilament and tubulin-rich neurites were found primarily on the lower side of the membrane, while soma, visualized with DAPI, were only present on the top on the membrane (Fig. 11&12). Neurites and soma were mechanically isolated from either side of the membrane for proteomic and transcriptomic analyses. As neurites are far smaller than soma and contain much less RNA and protein, it would be uninformative to compare the proteomes of neurite and soma fractions from the same number of cells. Instead, following recent approaches [81], [282], [295], in assumption of uniform total protein and RNA concentrations over the cell body the numbers were normalized according to the total protein and RNA content in the sample. All protein or RNA measurements in this study are therefore relative differences in the subcellular proteome or transcriptome.

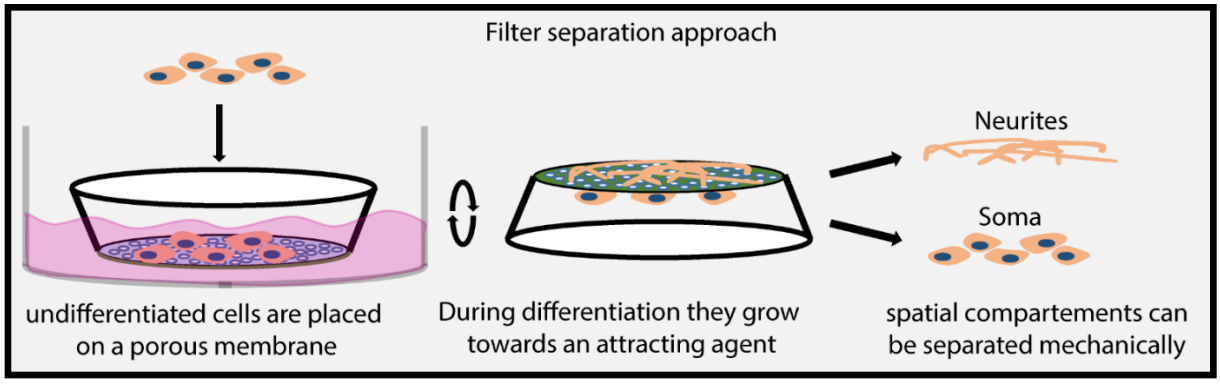


Fig. 10 Schematic presentation of the separation of subcellular compartments.

Separation approaches are always performed starting with 1M cells on a 6-well-insert with 4ml of medium in total. Mean yield of the soma fraction is about 15µg of RNA and 400µg of protein for N1E-115 cells and 1µg of RNA and 100µg of protein per filter for Ascl1 cells. Mean yield for the protrusion fraction is about 100ng of RNA and 15µg of protein per filter for N1E-115 cells and 200ng of RNA and 20µg of protein per filter from Ascl1 cells.

5.3.1. Local -omics analysis of the neuroblastoma cell line N1E-115

Staining differentiated cells on the membrane using markers for tubulin and DAPI confirmed a clean separation of cellular outgrowths and soma compartments (Fig. 11a). The average yield of protein extraction from one filter was 15µg per neurite sample and 400µg per soma sample.

Coomassie staining already showed the difference in protein composition between the samples (Fig. 11). Subsequent label-free protein quantification via mass spectrometry revealed a diverse proteome of the subcellular compartments. GO analysis of proteins enriched in the outgrowth showed a clear enrichment for membrane, cell-junction and receptor associated proteins (Fig. 11f). Marker proteins as histones (e.g. Histone 2A, Histone 2B) and components of the nuclear envelope (e.g. Lamin B1 and several nucleoporins) are found either exclusively or highly enriched in the somatic fraction. On the other hand, marker proteins associated with neurite growth cones (e.g. Netrin receptors, Slit2, Sema6d) and membrane protrusions (e.g. Actn4, Fmnl1) were found highly enriched in the neurite fraction. For validation we compared our dataset with a published proteomic dataset from Pertz et al. [282].

The extracted RNA was depleted from ribosomal RNA using RiboZero bead purification and used for total strand specific RNA-seq sequencing. In total, 19,998 genes could be identified by at least two readings in either of the compartments. Eight transcripts could be validated as upregulated more than twofold with a p-val of below 0.05, whereas 287 transcripts could be identified as being significantly enriched in the soma fraction more than twofold, with a p-val of below 0.05. Furthermore, qRT-PCR analysis confirmed that the dataset is statistically weak but holds true for these cells (Fig. 11e). Problems in the depletion of ribosomal RNA in one neurite sample resulted in a very shallow dataset for this replicate. This is the reason for the low number of enriched found transcripts and low statistical validity of the experiment.

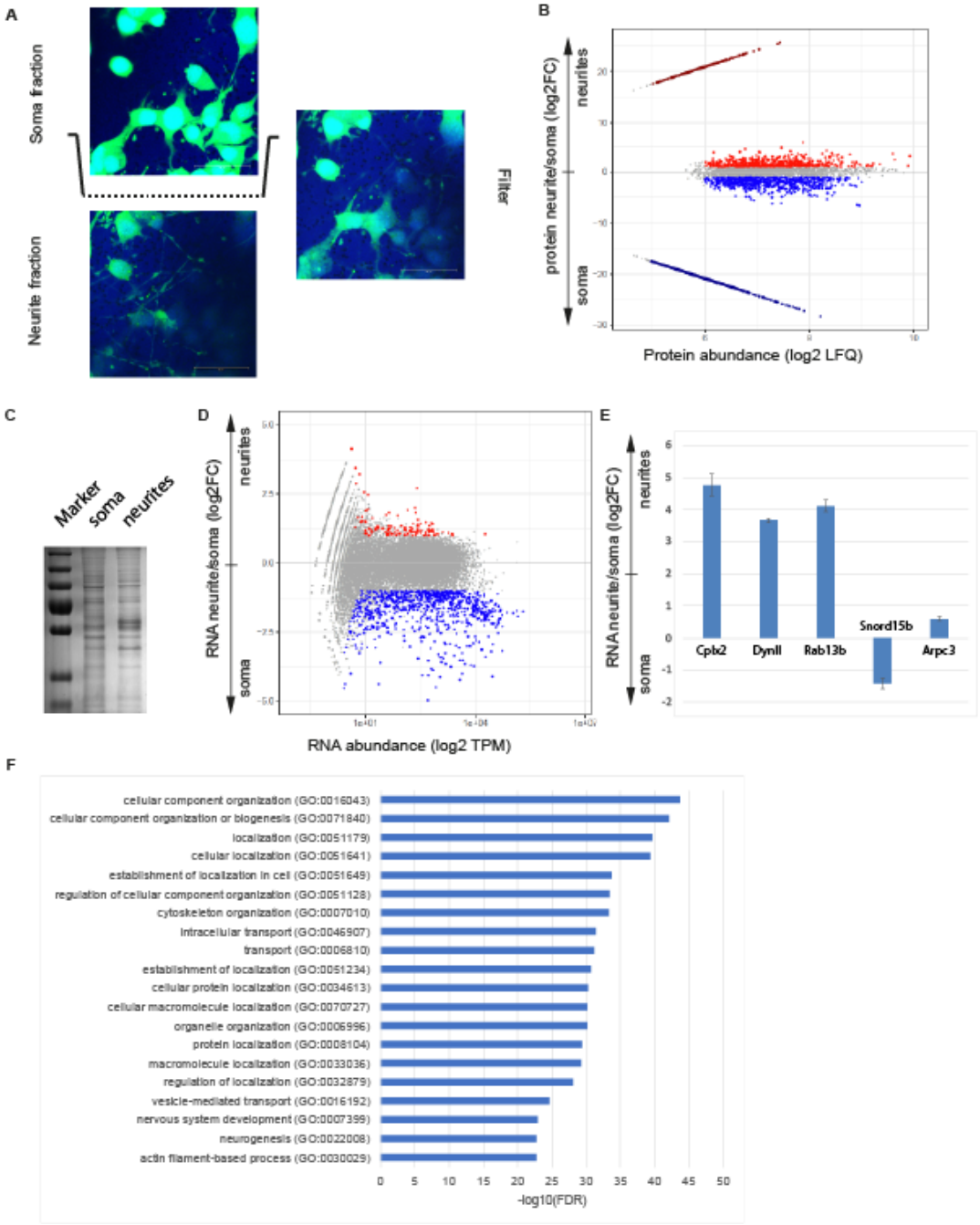


Fig. 11 Subcellular compartment separation of the neuroblastoma cell line N1E-115

A Immunofluorescence picture of differentiated cell on the membrane. Tubulin (green), DAPI (blue). Outgrowths below the filter are difficult to see because of autofluorescence signal coming from the PET-membrane. **B** Label-free MS of subcellular compartments of N1E-115 cells. Highlighted are protein that are exclusively found in the outgrowths (dark red), outgrowth enriched proteins (red), soma enriched proteins (blue) and proteins that are exclusively found in the soma fraction (dark blue). Primary analysis of the data was performed by Guido Mastrobuoni. **C** Coomassie staining of subcellular compartments from N1E-115 cells. The different patterns of protein compositions are clearly visible. **D** RiboZero total RNA-seq of RNA coming from subcellular compartments. Highlighted are transcripts enriched in the outgrowth (red) and transcripts that are enriched in the soma fraction (blue) with a p-val below 0.05. Primary analysis of the data was performed by Vedran Franke. **E**

qRT-PCR validation of some known localized transcripts. **F** GO analysis for the neurite enriched proteins and mRNAs.

5.3.2. Local –omics analysis of Ascl1 induced neurons derived from mouse embryonic stem cells

The clean separation of neurite and somatic subcellular compartments was validated via immunofluorescence imaging using markers for tubulin and neurofilament (Fig. 12a). The yield of protein extraction from subcellular compartments was 20µg per neurite fraction and 200µg per soma fraction.

The contamination of proteins coming from the coating reagent was prohibited using a SILAC approach. Unfortunately, the correlation of forward and reverse experiment was not sufficient for further analysis (Fig. 12b). Plotting the results of the label swap experiments against each other revealed a clear correlation but also a very shallow forward dataset. Since this might indicate a contamination of the forward neurite sample with soma material, the experiment was validated using a third dataset produced by lab colleague Alessandra Zappulo. As expected, the reverse experiment showed a clear correlation with the label-free quantification and is therefore used throughout the study (Fig. 12c&d). In total 4,317 proteins where identifiedby at least two unique peptides in the soma and the neurite sample. Since relative quantification by SILAC relies on a ratio of peptides, proteins identified exclusively in one compartment are not included in this analysis. 234 proteins could be identified as enriched more than twofold in the neurite fraction, while 179 proteins could be identified as enriched more than twofold in the soma sample. Western blot analysis of different marker proteins chosen on functionality and antibody availability confirmed this data. GO analysis revealed a clear enrichment for molecular transport and cytoskeleton in the neurites of Ascl1-iNs (Fig. 12e). Several genes known to regulate neurite outgrowth, cell protrusion and growth cones such as *Tnc* [296], *Fgf3* [297] or *Ptgs2* [298] showed a clear enrichment in the neurite sample. On the other hand, marker proteins for DNA metabolism, chromatin structure and the nuclear envelope are found highly enriched in the somatic fraction. All together this proves the clean separation and validates the SILAC experiment.

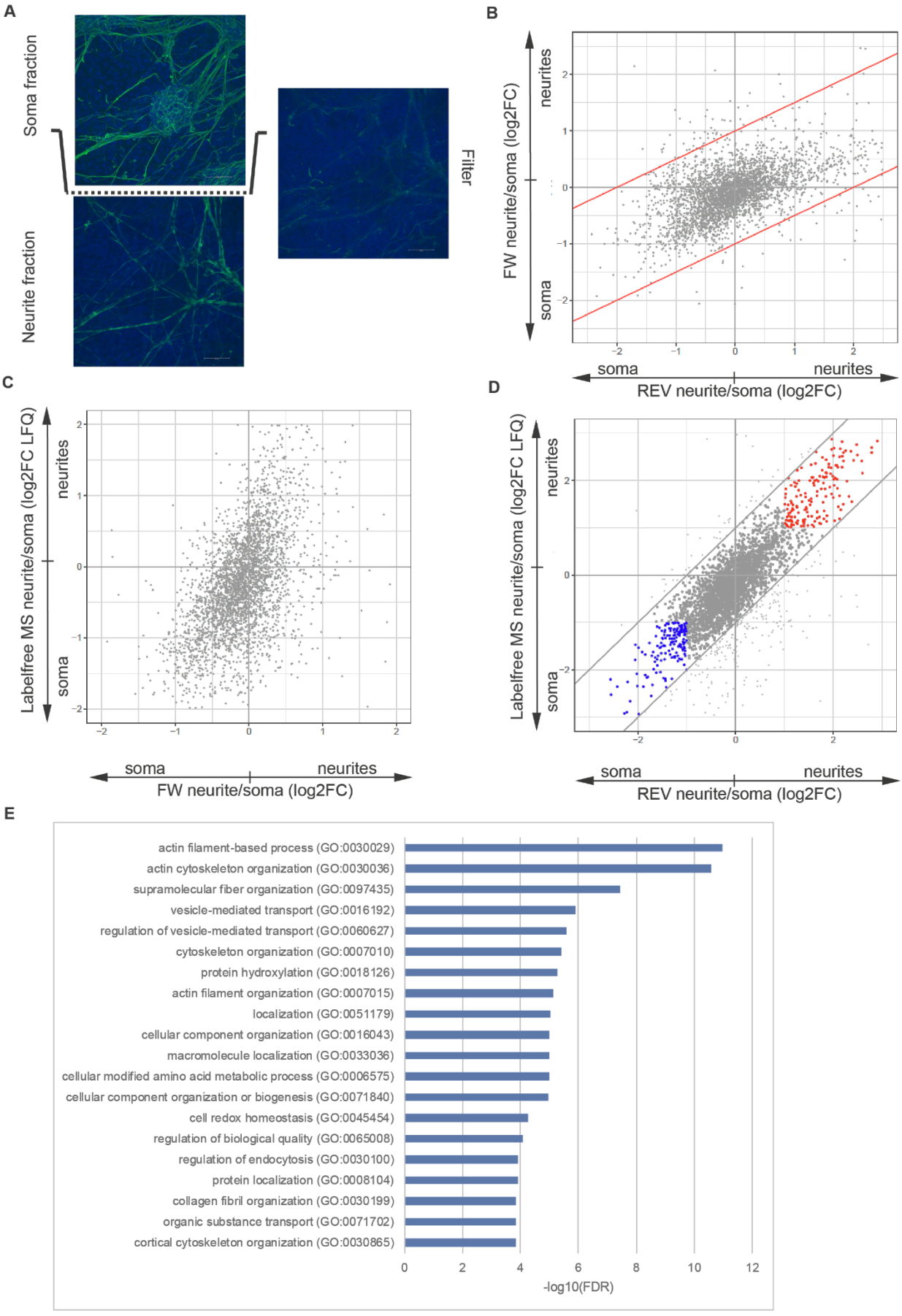


Fig. 12 Subcellular compartment separation followed by proteome analysis of Ascl-iNs

A Immunofluorescence of Ascl1-iNs on the filter. Pictures were taken five days after Ascl1 induction. Neurofilament (green) and DAPI (blue). The autofluorescence from the PET-membrane covers slightly the neuritic signal. **B** Subcellular SILAC of Ascl1-iNs. Differently labelled compartments are always compared. Label-swap replicates are not conclusive. Primary analysis was performed by Koshi Imami. **C&D** Validation of label swap replicated via label free quantification (done by Alessandra Zappulo). In C the forward replicate of the SILAC experiment is plotted against the label-free data. The correlation is not conclusive. In D the reverse experiment is plotted against the label-free data. The correlation is sufficient. Highlighted are neuritic enriched proteins (red) and somatic enriched proteins (blue). **E** Gene ontology analysis of enriched proteins.

The average yield for RNA extraction from Ascl1-iNs was 200ng per neurite fraction and 2µg per soma fraction.

After extraction and quality control the RNA was depleted of ribosomal RNA using RiboZero and submitted to a total stranded RNA-seq experiment. More than 18,000 protein coding genes in neurites and more than 19,000 protein coding genes in the somatic fraction were quantified with a threshold of one or more RPKM (reads per kilobase of transcript per million mapped reads).

In total 1,292 genes were identified as enriched in the neurites more than twofold with a p-val of below 0.05 and 3,614 genes were identified as enriched more than twofold in the soma fraction with a p-val of below 0.05 (Fig. 13a). The performed RNA-seq experiment was validated via nanoString and additionally 21 genes were validated by qRT-PCR (Fig. 13b & c). GO enrichment analysis revealed a clear overrepresentation of terms associated with neuronal functions, cytoskeleton, neuronal protrusions and growth cones in the neurites of Ascl1-iNs (Fig. 13d).

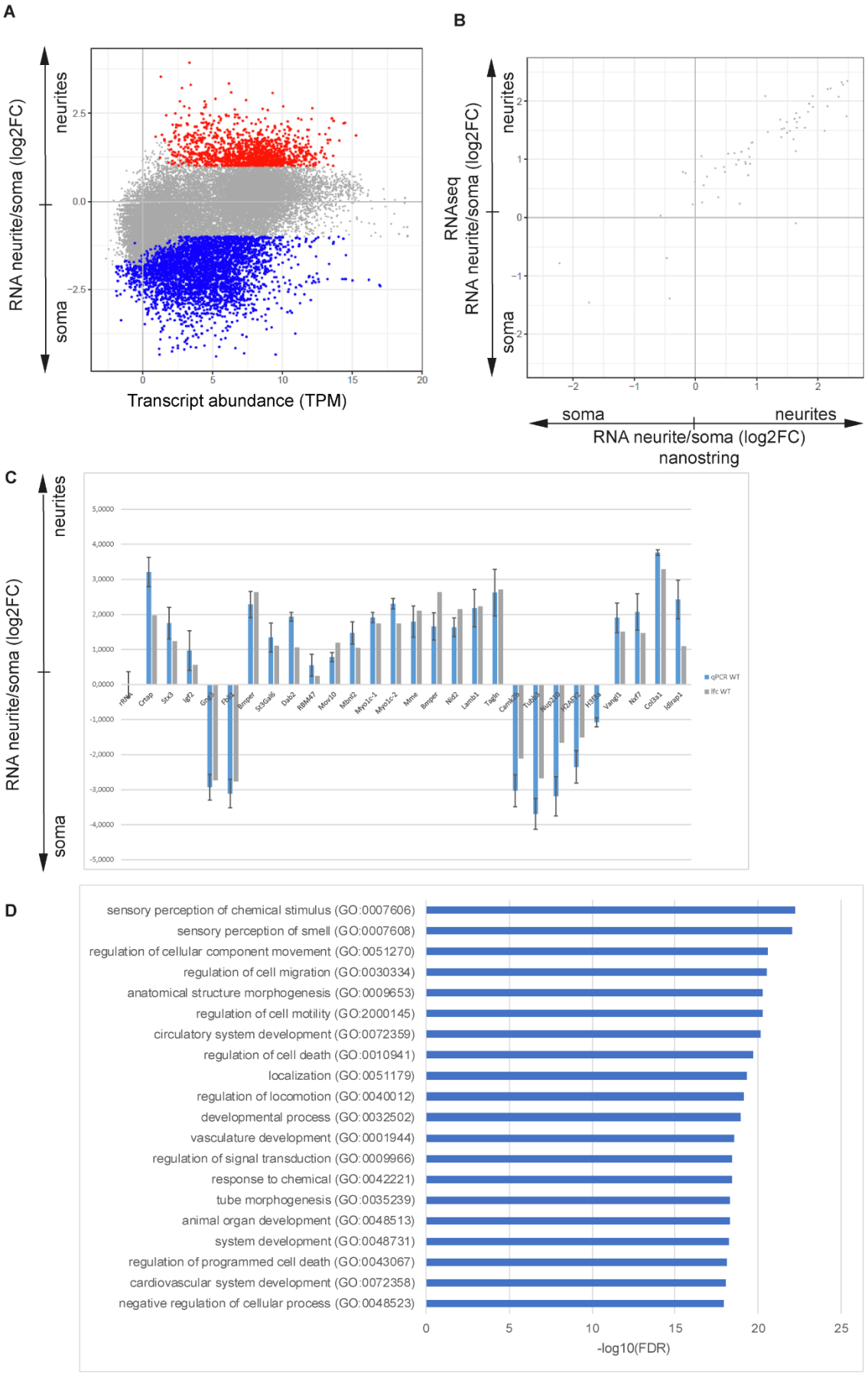


Fig. 13 Transcriptome analysis of subcellular compartments of Ascl1-iNs

A RiboZero total RNA-seq of subcellular compartments. Highlighted are neuritic enriched transcripts (red) and somatic enriched transcripts (blue) with a p-val below 0.05. Primary analysis was performed by Vedran Franke. **B** Validation of Transcript localization via nanoString. **C** Validation of RNA-seq via qRT-PCR. 21 representative genes were taken for a detailed validation of the RNA-seq results via qRT-PCR.

5.3.3. The local transcriptome and the local proteome is cell line specific

Previously published datasets showed that the spatial transcriptome and proteome are highly cell type specific [81], [109], [246], [249]. Comparison of the spatial enrichment of transcripts in neurite outgrowths of N1E-115 to neurites of Ascl1-iN resulted in very little overlap. A significant correlation between the relative enrichment measured in these experiments could not be observed (Pearson's r=0.07). No transcript shows enrichment in both cell lines by more than twofold and a p-val of below 0.05 and 16 transcripts show a somatic enrichment in both cell lines by more than 2-fold and a p-val of below 0.05. On the proteome level 61 proteins show a relative enrichment in both cell lines of more than twofold and 69 proteins show a relative somatic enrichment of more than 2-fold in both cell lines. However, no correlation between the measured relative enrichment in these cell lines could be observed for the proteome either (Pearson's r=-0,04).

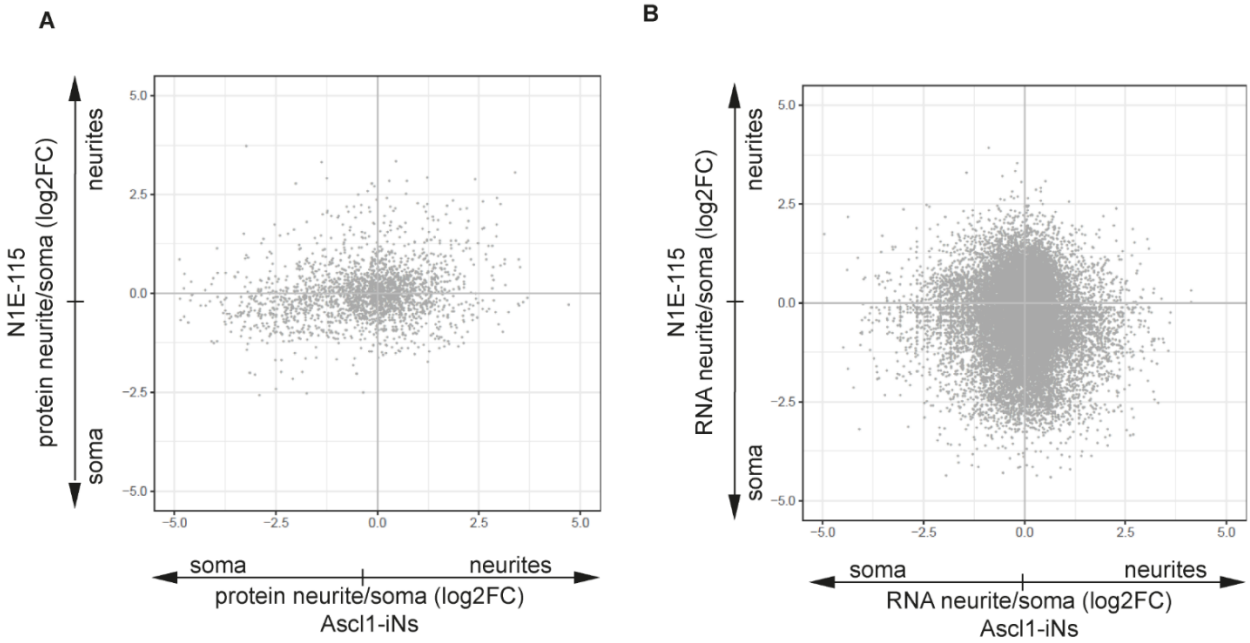


Fig. 14 Comparison of the neuronal model systems N1E-115 and Ascl1-iNs.

A Comparison of RNA localization between the two model cell lines. 34,256 transcripts could be identified by at least two reads in all compartments of both cell lines. Pearson's r=-0.05. **B** Comparison of protein localization between the two model systems. 3,269 transcripts could be identified with at least two peptides in all compartments of both cell lines. Pearson's r=0.075.

5.3.4. Identification of differential isoform localization, lncRNAs and circRNA localization

Differential splicing can be an effector of spatial RNA regulation and RNA localization [299] and previous studies highlighted the differential localization of isoforms for example in the tips of migrating fibroblasts and also in neuronal projections [81]. Regulatory sequences can be spliced differentially, and the stability of transcripts can be altered tremendously by alternative splicing. The identification of differentially spliced isoforms in total RNA sequencing data is not trivial due to complex splicing patterns. Employing MISO [111], 39 triplets of exons with a slightly different localization were identified (Fig. 15a). Alternatively, spliced exons are identified in this network by comparing the relative enrichment of one exon compared to its neighboring up and downstream exon. None of the identified exon specific localizations are higher than twofold. Specific experiments targeting alternative splicing or A-tailing of transcripts have been performed to analyze the impact of alternative splicing on spatial localization in more detail. Interestingly, most of the differentially enriched exons are part of circular spliced isoforms. Circular mRNAs are a special form of alternatively spliced mRNAs. They can inherit exons and introns of canonical transcripts and are defined by a head-to-tail splice junction, leading to a circular molecule. It was shown that circRNAs are abundant in neuronal cells and particularly in the synapse [300], [301]. Due to their circular nature they possess specific attributes such as resistance to exonucleases, lack of a cap and an A-tail structure and overlapping and sometimes even indefinite open reading frames. It was shown that they can be translated *in vivo* [302] or act as sponges for trans acting factors such as miRs [303], [304]. 362 genes with circular isoforms were identified in Ascl1-iNs with more than one RPKM spanning the circular splice junction in either compartments. Out of the 362 circRNAs, 32 showed two or more-fold enrichment in the neurite compartment while 64 showed two or more-fold enrichment in the somatic compartment. Moreover, for 90 circRNAs the circular isoform shows a significantly higher neurite localization than the linear form (Fig. 15b). Approximately half of these circular mRNAs show a neurite localization, while the linear isoform is not localized. This points towards an isoform specific mechanism of localization. To give an example, the circular transcript of Ephrin type-b receptor 2 (*Ephb2*), a receptor tyrosine kinase that functions in axon guidance [305], is more than ten times more abundant in neurites of Ascl1-iNs than its linear isoform. Whereas in the soma this ratio is reversed. Linear *Ephb2* is seven times more abundant in the soma than its circular isoform.

Different affinities to localization machinery may mediate the differential localization of linear and circular transcripts. We also cannot exclude the possibility of local splicing, especially as we find a number of splicing factors enriched in neurites.

Non-coding RNAs comprise a heterogenous and important group of genes with various roles in gene expression. Forty percent of lncRNAs show brain-specific expression patterns [306]; thus, lncRNA expression in the neurites and soma of iNeurons were analyzed. 550 annotated lncRNAs (> 10 RPKM) were detected. Since most lncRNAs are nuclear, it is not surprising that most lncRNAs show a somatic enrichment, although 12 lncRNAs of unknown function exhibited over twofold enrichment in neurites (P-values < 0.05). This result suggests they could contribute to neuronal polarity, spatial translation or act on or with local proteins.

A

gene name	chromosome	gene ID	bayes factor	differential localization
Zufsp	chr10	ENSMUSG00000039531	14.76	0.46
Fance	chr17	ENSMUSG00000007570	3.54	0.34
Uimc1	chr13	ENSMUSG00000025878	5.29	0.33
Slain2	chr5	ENSMUSG00000036087	3.46	0.31
Rims2	chr15	ENSMUSG00000037386	25.58	0.3
Acin1	chr14	ENSMUSG00000022185	11.81	0.3
Tipin	chr9	ENSMUSG00000032397	6.29	0.26
Apc	chr18	ENSMUSG00000005871	21.39	0.25
Huwe1	chrX	ENSMUSG00000025261	1E+12	0.24
Mdm1	chr10	ENSMUSG00000020212	2.87	0.23

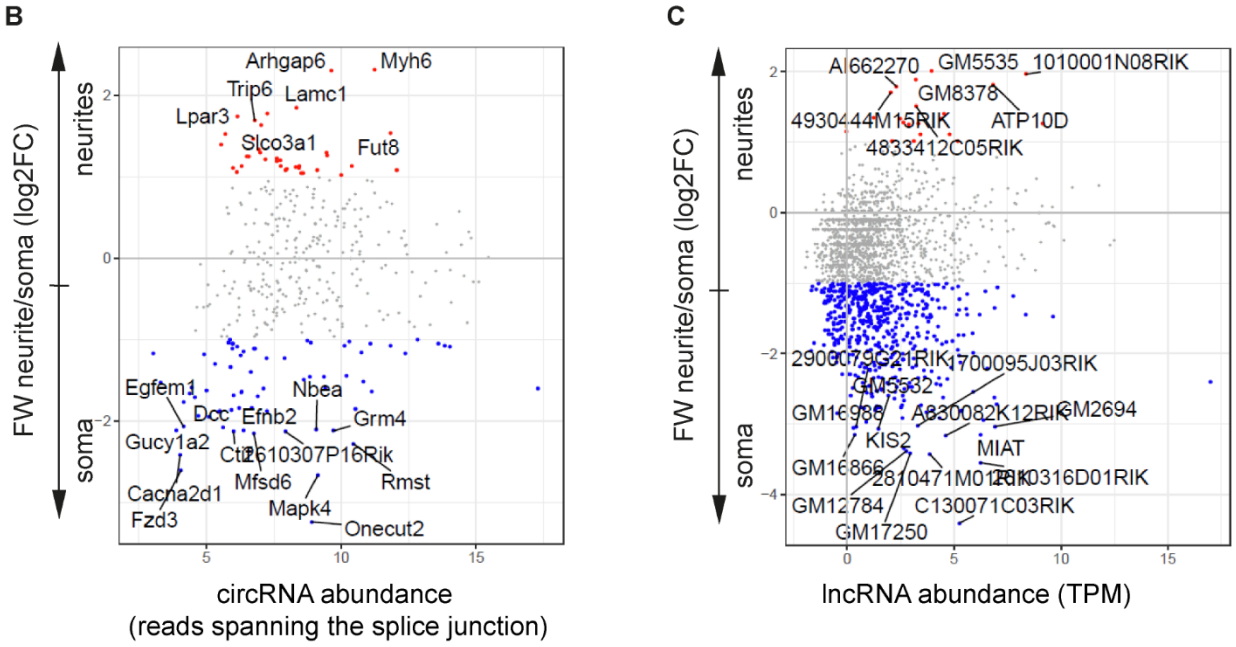


Fig. 15 Alternative splicing and localization.
A Differential exon localization in Ascl1-iNs. Primary analysis was performed by Vedran Franke. **B** Differential circRNA localization in Ascl1-iNs. Primary analysis was performed by Andrei Filipchyk. **C** Differential lncRNA localization in Ascl1-iNs. Primary analysis was performed by Andrei Filipchyk.

5.3.5. RNA localization correlates with protein localization in Ascl1 derived neurons

The impact of local protein synthesis on the spatial proteome has, despite its unquestionable importance, never been shown on a global scale. This phenomenon would imply a correlation of transcript and corresponding protein localization. To highlight the relation between transcript and protein product localization the datasets of subcellular protein and RNA localization were compared. While in N1E-115 cells the localization of transcripts and their corresponding protein products is not correlating (Pearson's r=-0.04) in Ascl1-iNs a clear and positive correlation is found (Pearson's r=0.55) (Fig. 16a&c).

This result might indicate that mRNA localization is important for a substantial fraction of the neurite-localized proteome.

Interestingly, the top enriched fraction of proteins in our dataset is also found spatially enriched as transcripts. This suggests that high amounts of local proteins require mRNA localization.

The average localization for transcripts for either neuritic or somatic enriched proteins does not differ in N1E-115 cells, while there is a significant difference in Ascl1-iNs ($p<0.005$; Fig. 16b&d). For further analysis the pool of genes was divided into six major groups: (1) genes for which the RNA and the protein were found enriched in the neurites; (2) genes for which the RNA and the protein product were found enriched in the somatic fraction; (3) genes for which the RNA was enriched in the neurites and the protein product was not; (4) genes for which the protein was found enriched in the neurites but the transcript was not; (5) genes for which the RNA was somatic enriched but the protein was not; and (6) genes for which the protein was somatic enriched but the RNA transcript was not. GO analysis revealed a clear enrichment for neurite and receptor associated genes within groups 1, 3 and 4 (Fig. 16e).

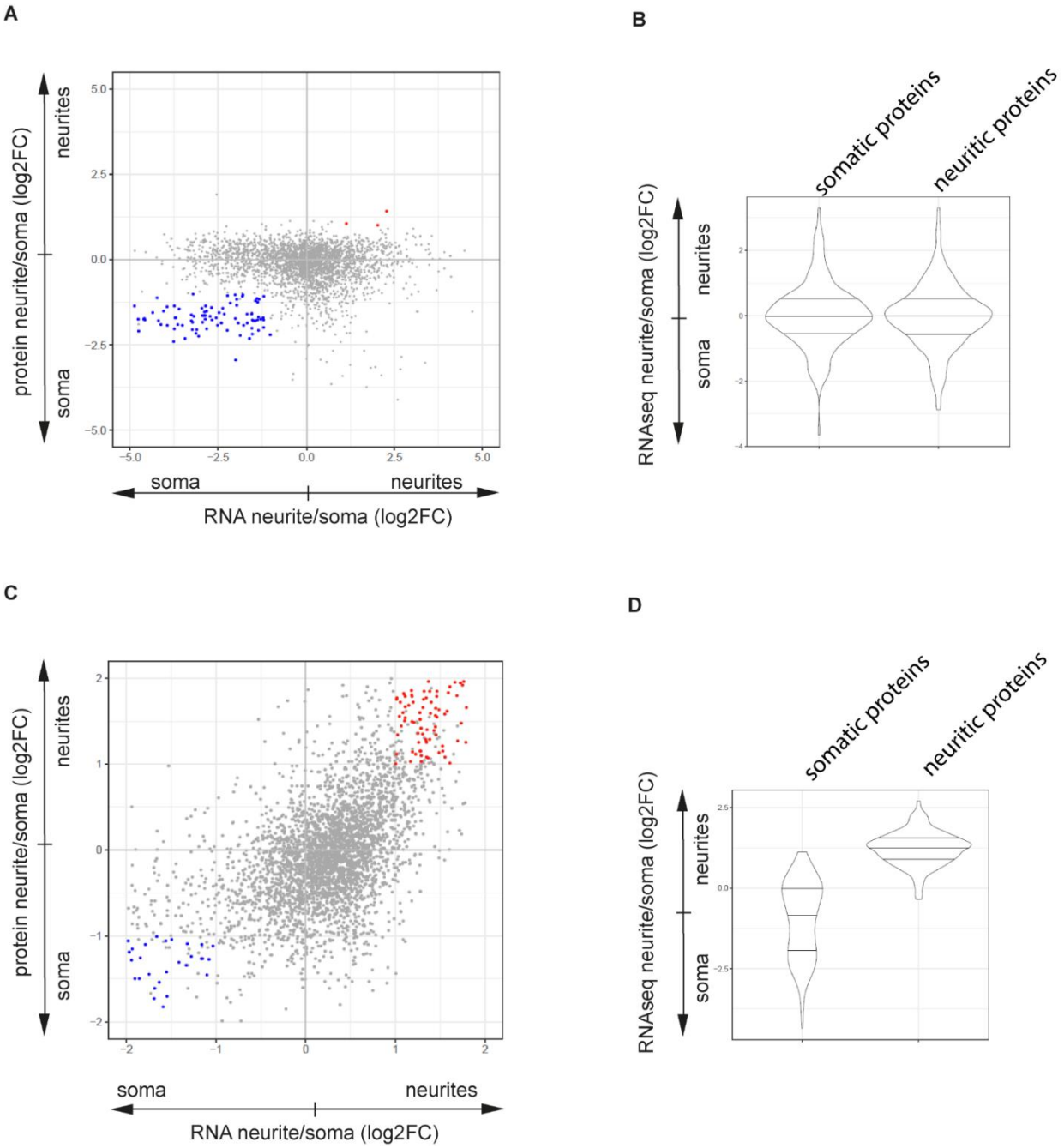


Fig. 16 Comparison of transcript localization and corresponding protein products in neuronal model systems

A Correlation of RNA-transcripts and protein products in the neuroblastoma cell line N1E-115 . Highlighted are genes for which transcript and protein are neurite localized (red) and genes for which transcript and protein are somatic localized (blue). **B** Violin plot for RNA localization of neuritic and somatic proteins. **C** Correlation of transcript and protein localization in Ascl1-iNs. Highlighted are proteins for which the transcript and the protein is neuritic localized (red) and genes for which the transcript and the protein product are somatically localized (blue). RNA and protein localization show a positive correlation (Pearson's r=0.55). **D** Violin plot for the transcript localization of either somatic or neuritic proteins. **E** Gene ontology analysis for enriched genes.

5.3.6. Overlay with spatial ribosome profiling data reveals the impact of local translation on the spatial proteome

To rule out the possibility of independent localization of proteins and their transcripts the spatial translatome was examined. The impact of local protein synthesis on protein localization can be visualized by correlating datasets for spatial proteome, transcriptome and translatome. For translatome, a dataset from spatial ribosome profiling (RiboSeq) experiments performed by Camilla Cioli Mattioli was used [307]. In these experiments just those mRNAs were sequenced which were in the moment of fixation covered by a ribosome. Assuming that the elongation state is predominant, this can be used as a proxy for ribosome activity and therefore translation. We used the ratio of RiboSeq reads in neurites versus soma to assess the relative translation amount in the spatial compartments (Fig. 17a). The ratio of RiboSeq reads per kb of transcript can therefore be taken as a measurement of translational efficiency of a specific transcript. The spatial RiboSeq data correlates with both protein (Pearson's r=0.41) and RNA localization (Pearson's r=0.33). Comparing the relative RiboSeq signals for the different subgroups shows a significant difference between somatic and neuritic enriched transcripts (Fig. 17b).

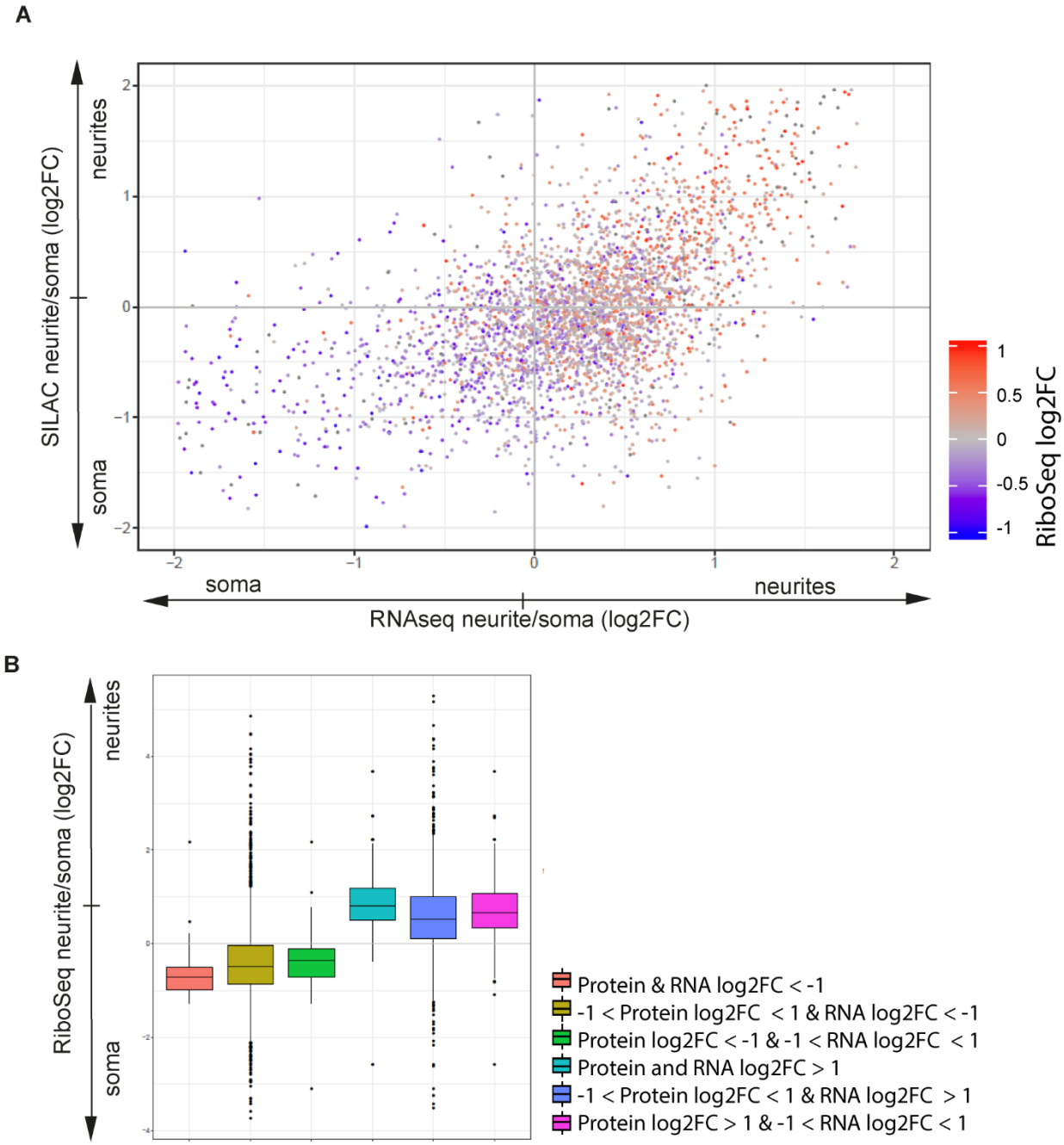


Fig. 17 Local translation in neurites of Ascl1-iNs

A Local translation identified via subcellular RiboSeq correlates well with RNA and protein localization. Plotted is RNA vs. protein localization color coded for local translation. RiboSeq data was generated by Camilla Cioli Mattioli. Primary analysis was performed by Lorenzo Calviello. .**B** Boxplot for relative RiboSeq signal for different gene subgroups.

For validation of these experiments spatial pSILAC and QuaNCAT experiments were performed. These experiments quantify *de novo* protein synthesis using pulse labelling of newly synthesized proteins via stable isotope labelled amino acids, followed by mass spectrometry. During a defined pulse, cells were treated with either heavy or medium labelled

amino acids that are subsequently incorporated into newly synthesized proteins. The pulse length is critical to distinguish between spatially synthesized and transported proteins. Due to the relatively low translational activity in post mitotic cells a pulse length of two hours was used in the pSILAC assay, knowing that protein transport in this time might bias the data. Ascl1-iNs cultured on a porous membrane were incubated for the pulse length in either heavy (H) or medium (M) isotope-labelled medium and subsequently used for subcellular fractionation of soma and neurites. Subsequently, the differentially labelled lysates were pooled (H neurites + M soma in forward (FW) and M neurites + H soma in reverse (REV) experiment, Fig. 18a). The FW and REV experiments represent "label swap" biological replicates to eliminate biases introduced by the labeling procedure (Fig. 18b). The ratios of peak intensities, H /M in FW experiment and M/ H in REV experiment, quantifies the relative translation rates in neurites versus soma. Thus, the translation rates of 242 proteins were measured in the two compartments (Fig. 18c, the relatively low coverage is expected after a short two hour labeling pulse). Importantly, we observed a strong correlation between relative translation rates measured by RiboSeq and pSILAC (Fig. 18c).

To shorten the pulse length even more a QuaNCAT experiment was performed to quantify relative translation rates in neurites and soma. These experiments combine isotope labelling of *de novo* synthesized peptides with azidohomoalanine (AHA; Fig. 18e) labelling. Azidohomoalanine enables covalent linking of newly synthesized proteins to alkyne agarose beads using "click chemistry", and their separation from the protein pool. Proteins are digested "on bead" and treated using the pSILAC workflow (Fig. 18e&f). The purification step employed in QuaNCAT substantially reduces the background of pre-existing proteins, which enabled reproducible measuring of relative protein abundance of 380 newly synthesized proteins after a short pulse of AHA (30 min). Relative translation rates measured by QuaNCAT and pSILAC supported the spatial omics analysis (pSILAC Pearson's r=0.53, QuaNCAT Pearson's r=0.71) and correlate positively to the RiboSeq dataset (Pearson's r=0.8), indicating that the pulse length and therefore protein transport is in this case not interfering with results.

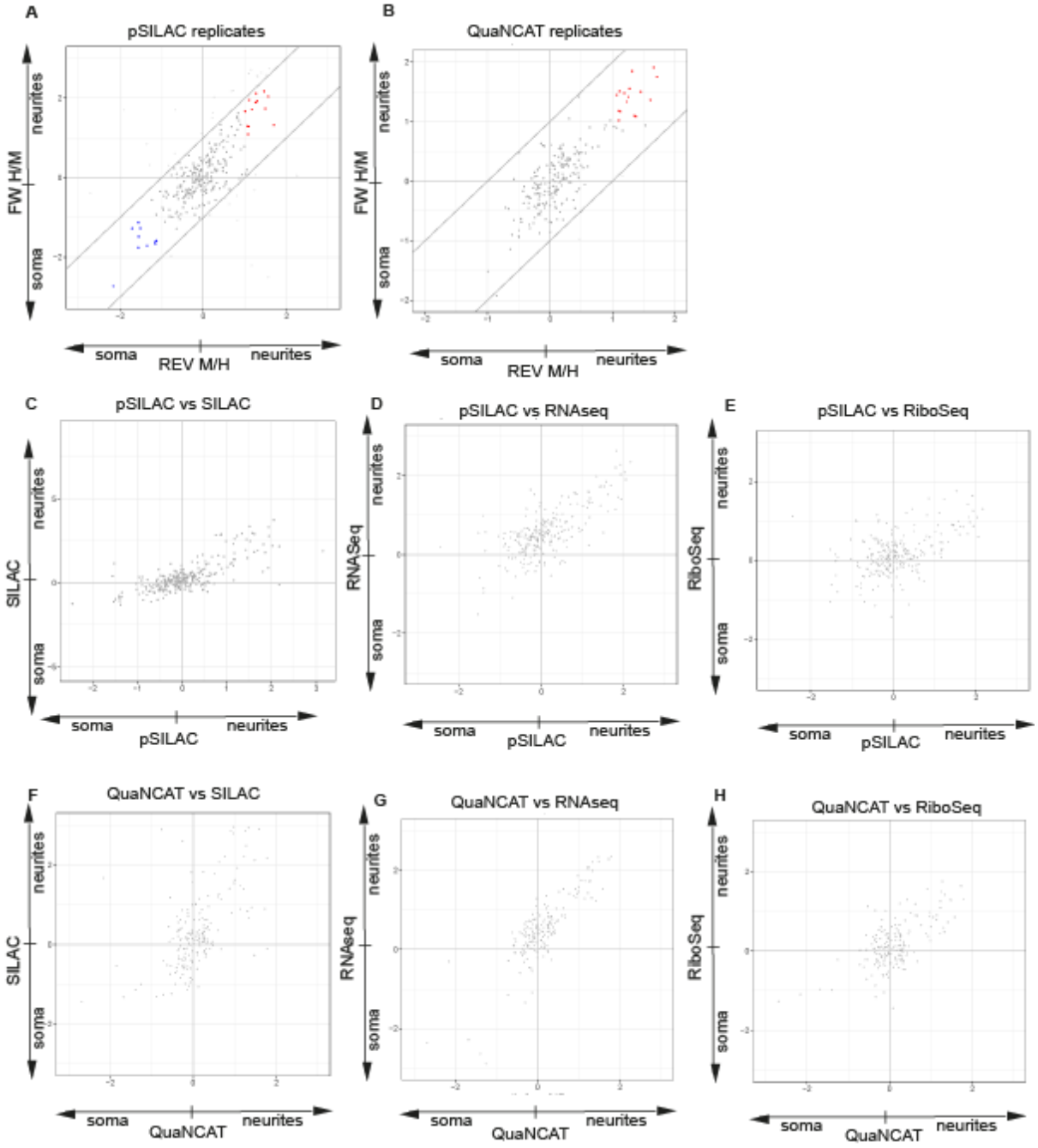


Fig. 18 Validation of local translation

A Scheme of pSILAC experiments. Cells were cultured in light medium for five passages prior to the experiment. **B** Label-swap replicates show a sufficient correlation. Outliers were excluded from further analysis. Primary analysis was performed by Koshi Imami. **C** Correlation of pSILAC and RNA-seq results. **D** Correlation of pSILAC and SILAC results. **E** Scheme of QuaNCAT experiments. Cells were cultured in light medium for five passages prior to use. **F** Label-swap replicates show a sufficient correlation. Outliers were excluded from further analysis. Primary analysis was performed by Erik McShane. **G** QuaNCAT results plotted against local RNA-seq results. **H** QuaNCAT results potted against SILAC results.

5.3.7. Pools of localized RBPs and miRs as potential players of RNA localization and regulators of local translation

Local translation in neurites is very likely to be controlled as in any other part of the cell. Major players of posttranscriptional regulation are miRNAs and RBPs. In addition to that, RBPs and miRNAs bare the potential to play a role in the process of localization itself. To assess the number of miRNAs in the spatial transcriptome, small RNA sequencing experiments were performed. Hereby, small cDNAs were enriched through gel purification prior to sequencing. 500ng input material coming from either outgrowth or soma of N1E-115 or Ascl1-iNs was used as input material. In total 933 miRNAs in N1E-115 cells were identified, of which 20 are more than twofold enriched in the outgrowth and 14 are more than twofold enriched in the soma fraction (Fig. 19a). Among the outgrowth enriched miRNAs are miR-140-5p, which has a postulated neuroprotective function [308] and miR-5115, which is known to be localized to the synapse [309].

In Ascl1-iNs 1,648 miRNAs could be identified, of which 191 were more than twofold enriched in the neurites and 151 more than twofold enriched in the soma (Fig. 19b). Among the top neurite enriched miRNAs there are several for which neuronal functions were already shown,such as miR-214-3p and -5p [310], miR-483-3p and -5p [308] and members of the let-7 family [311], [312]. No correlation in miRNA localization between the cell lines could be detected (Pearson's r=-0.01). Components of the RISC were found in the neurite compartments of both cell lines but were not enriched in either of them.

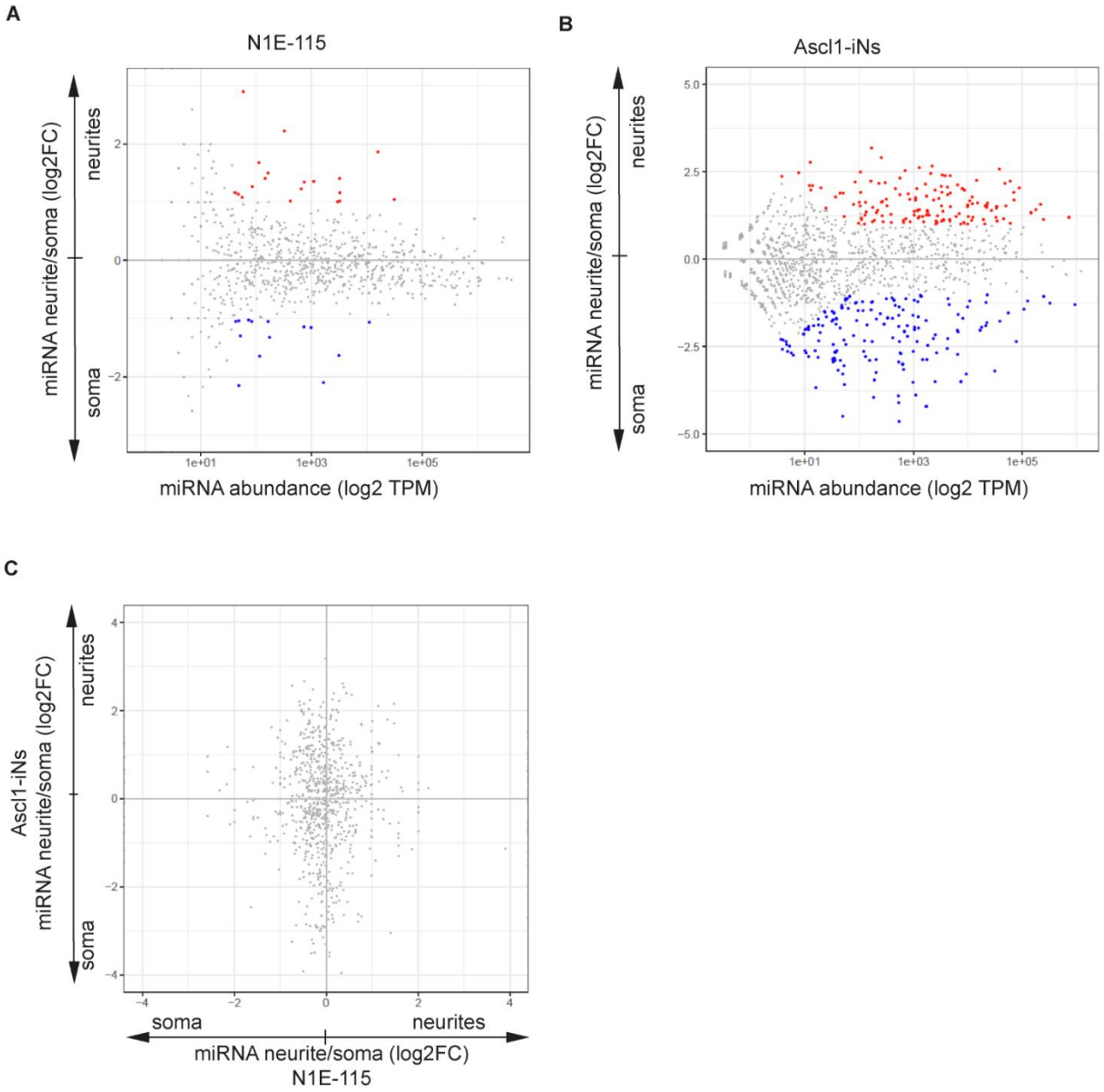


Fig. 19 Spatial miRs in the neuronal model systems N1E-115 and Ascl1-iNs

A Spatial small RNA-seq reveals local miRNAs in N1E-115 cell. 933 miRs were identified of which 20 are enriched in the outgrowth (red) and 14 miRs are enriched in the soma (blue) fraction are. **B** Spatial small RNA-seq reveals local miRNAs in Ascl1-iNs. 1,648 miRs were identified of which 191 miRs enriched in the neurites (red) and 151 miRs enriched in the soma (blue) are. Primary data analysis performed by Vedran Franke. **C** Correlation of miR localization between the used cell types.

Targets of identified miRNAs were chosen using a dataset published by Chou et al. 2017 [313]. Only experimentally validated miRNA-mRNA interactions were used for the analysis. As parameters for validated miRNA-target interactions were chosen pairs of miRNA-mRNAs identified by qRT-PCR, RNA-Seq, microarray, luciferase reporter system and western blot respectively. Localization of miRNAs does not correlate with their validated targets, in N1E-115 not in Ascl1-iNs (Fig. 20a&b). Also, there is no observable correlation between miRNA

localization and differential spatial translation efficiency of the target mRNA in Ascl1-iNs (Fig. 20b).

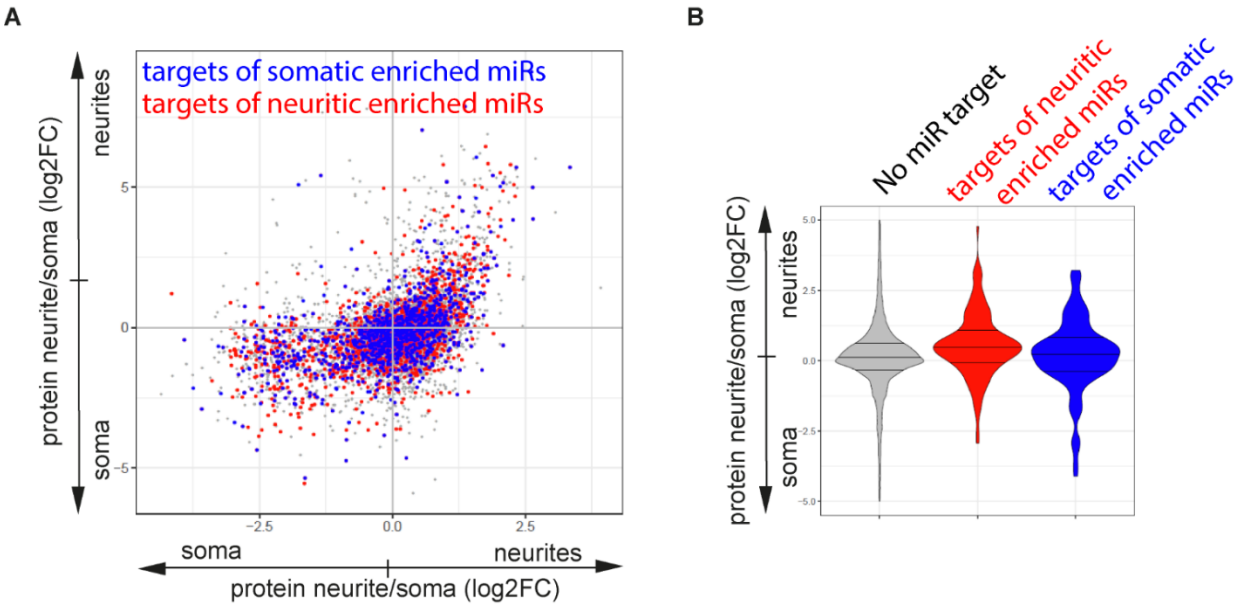


Fig. 20 Targetome of either neuritic or somatic enriched miRs in Ascl1-iNs

A Localization of miRs and their targets does not correlate in Ascl1-iNs. Targets of neurite enriched miRs highlighted in red, targets of somatically enriched miRs highlighted in blue. **B** Boxplot of spatial relative translation efficiency of validated miR targets. There is no significant difference of the spatial relative translation efficiency of targets of neuritic enriched miRNAs or somatically enriched miRNAs.

Recent studies identified mRNA-bound RBPs in diverse cells using an mRNA interactome capture approach and created a census of 1,542 RBPs [314]. Overlaying spatial proteome data from the label-free quantification approach in N1E-115 cells and the SILAC based experiment in Ascl1-iNs with said RBP dataset generated by Tuschl et al. reveals the distribution of RBPs in N1E-115 and in Ascl1-iNs (Fig. 21a & b). In total, 730 RBPs could be identified in N1E-115 cells by two or more peptides, of which 12 were found more than twofold enriched and six exclusively in the outgrowth. On the other hand, 476 were found more than twofold enriched in the soma and 119 exclusively in the soma. In Ascl1-iNs 806 RBPs were identified as being highly expressed, of which 26 were found more than twofold enriched in the neurites and 52 were more than wofold enriched in the soma. Neurite enriched RBPs fall into diverse categories: ribosomal proteins (Rpl26, Rpb27, Rpl31) and rRNA modulators (Urb2, Tfb1m), translational regulators (Lsm14b, Rrbp1), proteins involved in tRNA metabolism (Wars, Rpp14) but also transcription factors, and proteins involved in DNA binding (Peg10, Ttf2, Pura, Gtf2f1), nucleases (Hrsp12, Rexo2), RBPs that are known to play a role in the localization

machinery (Mvp, Rrbp1, Elavl2, Lsg1, Nxf7) and a group of proteins with general RNA binding properties (Mov10, Calr, Rbms2, Ppan, Hdlbp, Fam120a, Rnaseh2c). Interestingly the average distribution of RBPs in the two examined cell lines is significantly different. Whereas RBPs in general are evenly distributed in Ascl1-iNs, in N1E-115 cells they are clearly somatically enriched (Fig. 21c).

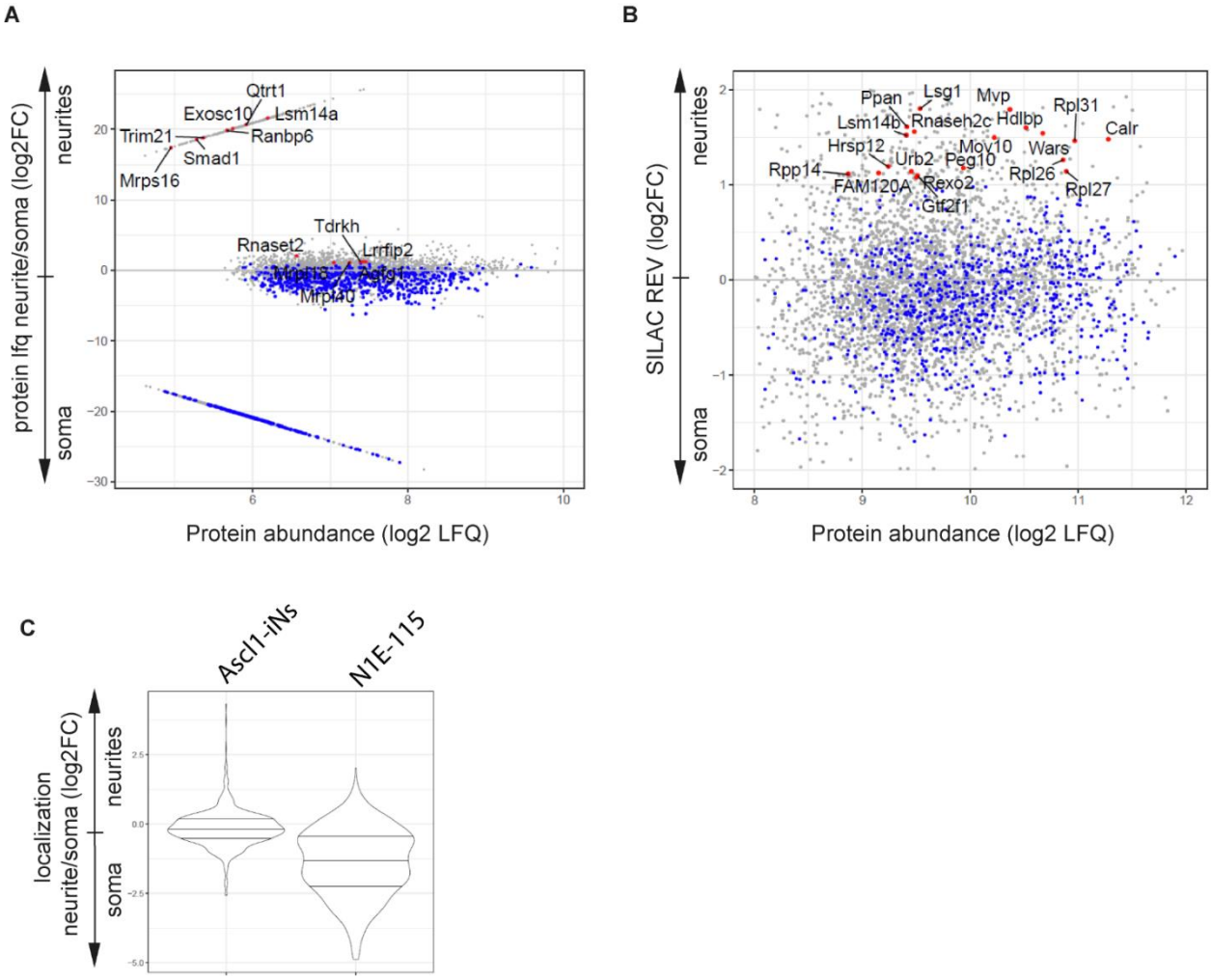


Fig. 21 RBP localization in the used model systems

A Overlay of the RBP dataset published by Tuschl et al. [314] and protein localization in N1E-115. Highlighted are RBPs enriched in the outgrowth (red) and other RBPs (blue). **B** Overlay of the RBP dataset with protein localization in Ascl1-iNs. Highlighted are RBPs enriched in the neurites (red) and not enriched RBPs (blue). **C** Violin plot for relative enrichment of RBPs in the used cell lines.

5.4. Going postal: Towards a zip-code directory for neuronal cells

5.4.1. Motive enrichment analysis of neurite enriched RNAs

There are various ways in which RBPs can target mRNAs: based on linear motifs, secondary structures, unstructured regions, or even unspecific targeting the ribose backbone. Most of these interactions occur in the untranslated regions of the mRNA, thus either the 3' or 5' UTR. In order to identify *de novo* sequence motifs for mRNA localization the MEME [315] Suite toolbox and MAST [316] were employed to identify motifs enriched in the UTR of localized transcripts. As a background pool, somatically enriched or equally distributed mRNAs were used. We found several motifs as significantly enriched in the neurites of Ascl1-iNs compared to the soma (Fig. 22). Some of them are known to be targeted by RBPs or other mechanisms. For example, 21% of all mRNAs that are neurite localized and also spatially translated showed a reminiscent GC-motif that is associated with m1A methylation sites in the 5' UTR region of mRNAs and that is reported to promote translation [317]. Overall, reported m1A sites from mouse liver, mouse fibroblasts, mESC and brain [318] showed a significant overlap with the neurite transcriptome. Some of the other *de novo* found motifs match to known RBP binding motifs, such as the hnRNP E/poly(rC)-binding protein Pcbp2, which is broadly involved in RNA metabolism. It was shown that Pcbp2 regulates splicing for *Mapt/Tau* exon 10, which is critical for neuronal survival and function [319].

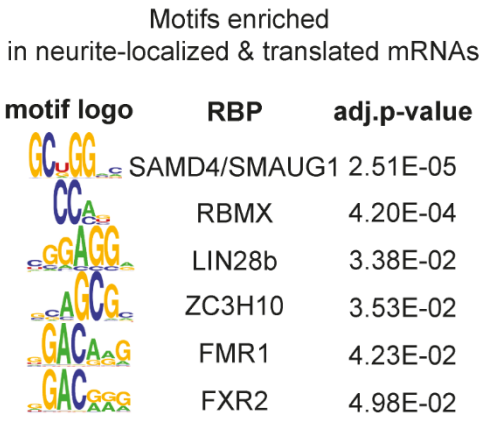


Fig. 22 Motive enrichment analysis of localized Transcripts.

A Enriched motifs in neurites of Ascl1-iNs identified by MAST and MEME. Primary data analysis performed by Esteban Peguero Sanchez.

5.4.2. Using an RNA reporter library to identify zip-codes in Ascl1 iNs

The high throughput identification of regulatory RNA sequences is a wide and fast-growing field in molecular biology. Several techniques are known and widely used to identify the interactions of specific RBPs with RNAs (such as RIP [320] or CLIP [321]) or to find regulatory elements inside restricted sequences. Here a technique is proposed to identify RNA-elements *de novo* that are targeted by the localization machinery. Although the method can discern sequence elements which are necessary for localization, it can discriminate between the different mechanisms of localization. The theory behind these experiments is to generate an RNA-reporter library containing a pool of overlapping RNA constructs, deliver this pool into neuronal cells and recover the reporter pool from subcellular compartments after a period of incubation. Differences between the pools recovered from neurites and soma can be used to identify (1) sequences that are actively localized to one of the compartments and (2) sequences that vary in stability between the compartments (Fig. 23). To ensure highest sensitivity in reporter readout, next generation sequencing methods were used to validate the input and output of the experiments.

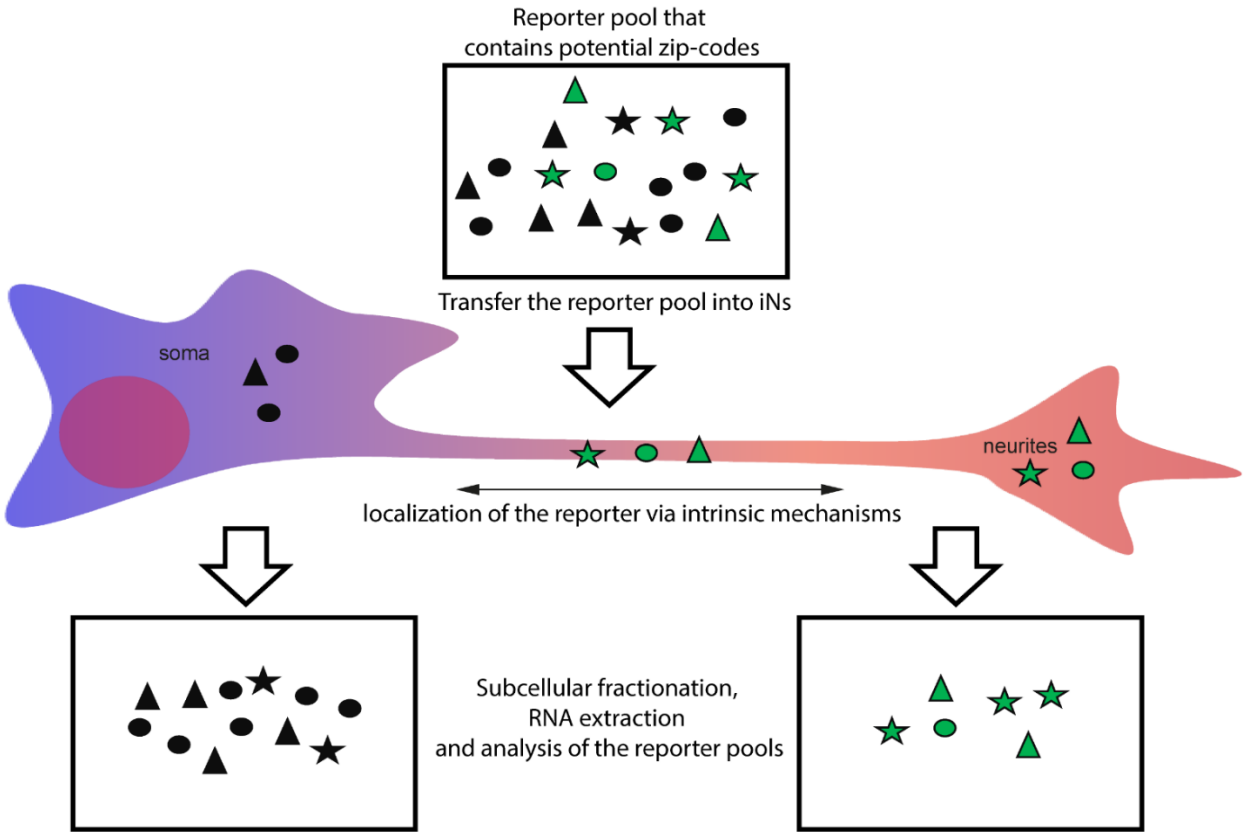


Fig. 23 Schematic presentation of high throughput identification of zip codes in Ascl1-iNs.

Reporter libraries are delivered into Ascl1-ins and the localization of these reporters is analyzed using NGS techniques. Experiments are always performed as batch experiments to guarantee reproducible results.

The input reporter pool of transcripts has to fulfill several criteria in order to be useful for the identification of zip-codes in neurons: (1) regulatory sequences have to be in the pool but the background should be non-regulatory; (2) sequences have to be long enough to contain full sequences yet short enough to enable specific discrimination of the sequence motif; (3) the sequences have to be partially overlapping and redundant, to ensure tiling and therefore precise mapping of regulatory elements; (4) the reporter pool should be as diverse as possible; (5) the reporter has to mimic a nascent RNA; and (6) the reporter has to be easily and clearly distinguishable from native RNAs. To fulfil all of these criteria the input sequences for the first experiments were taken from a RiboZero total RNA Sequencing library coming from the neurites of Ascl1-iNs. These fragments fulfill most of the above-mentioned criteria: (1) they contain endogenous sequences; (2) the average size of the endogenous fragments is 180bp; (3) sequences are redundant and overlapping due to the nature of RNA-seq library preparation; (4) it contains, in theory, all possible sequences coming from the neurite compartment; (5) after in vitro transcription and A-tailing the sequences mimic endogenous RNAs; and (6) due to adapter sequences that are added during library preparation, the reporter pool can be easily separated from endogenous RNAs.

The RNA-seq library was amplified by PCR using primer overhangs that added a T7 promotor sequence and subsequently reverse transcribed them using a T7 in vitro transcription kit. To increase stability of the reporter pool the fragments were A-tailed to the degree of a nascent transcript (Fig. 24a). This RNA pool was introduced via RNA transfection into fully differentiated Ascl1-iNs cultured on porous membranes. After one day of incubation total RNA was extracted from the subcellular compartments and the reporter pool coming from the different compartments was isolated, making use of the sequencing adapters (Fig. 24a). The pool was sequenced on the NextSeq500, running one sample per flow cell with a read length of 150bp. Libraries were sequenced and mapped to the mm9-genome. 75,519,699 reads could be uniquely mapped to the genome in the neurite sample and 82,845,155 reads could be uniquely mapped to the genome in the soma sample. The average read distribution over mRNA molecules showed no overall bias towards any part of the mRNA. Amplification bias during the preparation of the library resulted in an extreme inhomogeneity of the signal, making the use of conventional peak calling or differential expression analysis uninformative.

A 10-mer “counting” approach was used to identify short linear motifs that are either enriched or depleted in one of the compartments. The distribution of all possible (4^{10} = 1,048,576) polynucleotides 10 bases long was calculated within the spatial compartments. Out of the 1,048,576 10-mers, 994 showed at least a twofold enrichment in the neurite sample.

Additionally, just those 10-mers with more than 100 reads were filtered out leaving a pool of 43 10-mers for the neurite sample and 83 10-mers more than twofold enriched in the soma sample. Plotting the neurite-enriched 10-mers to the genome shows a random distribution (Fig. 24b).

The top 10 enriched 10-mers for both soma and neurites were cloned in a row into the context of an artificial 3'UTR of a GFP reporter construct. This construct was stably integrated into the genome of Ascl1-mESC and localization was analyzed after differentiation and compartment separation via qRT-PCR (Fig. 24c). The soma reporter showed around fourfold somatic localization, while the neurite reporter showed just a slight somatic localization.

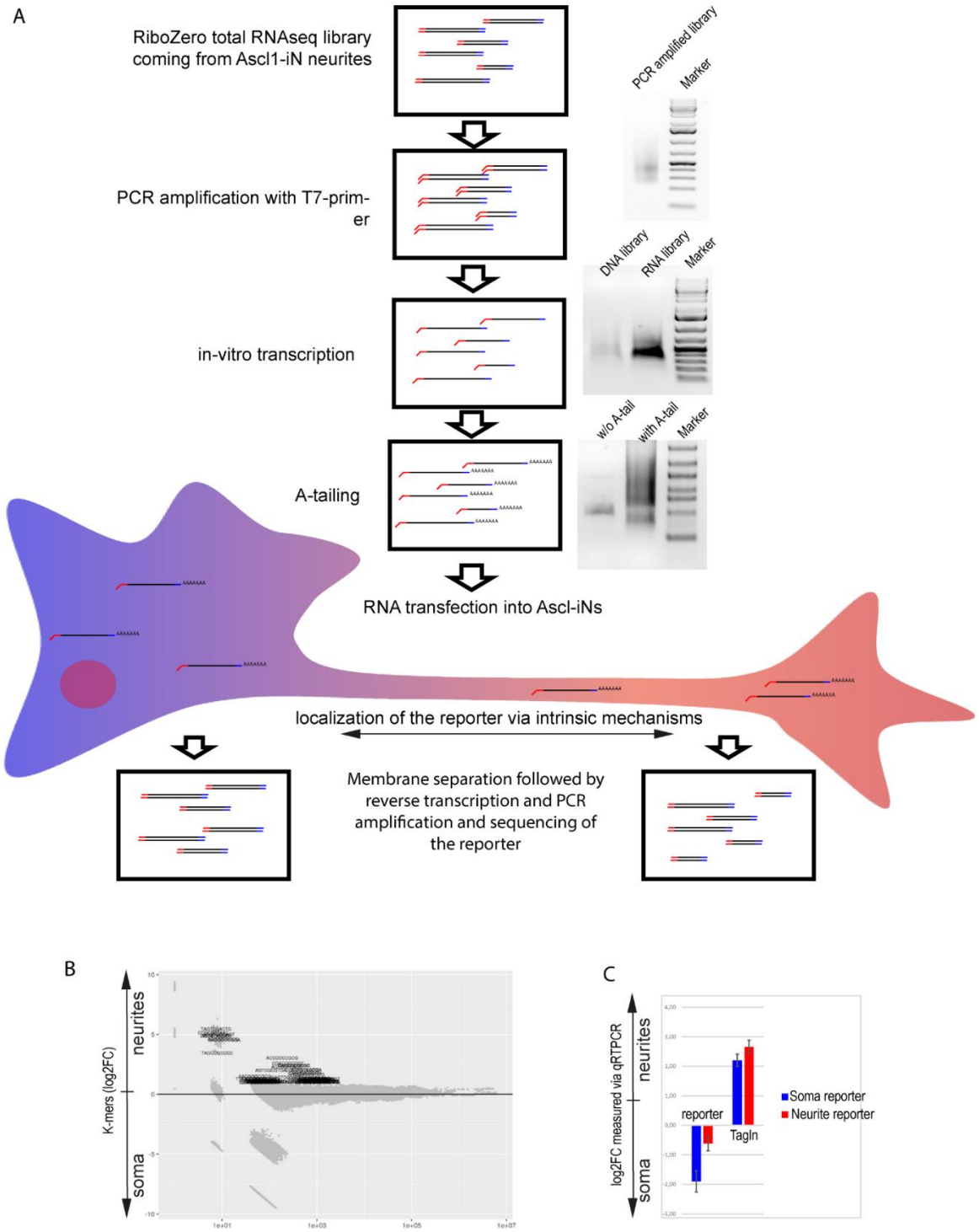


Fig. 24 Employing a total and transfected RNA reporter library in Ascl1-iNs.

A Experimental setup. Batch transfection of the highly diverse reporter library was performed using 300µg of RNA reporter library into 3,000,000 cells. All steps of the protocol were controlled using agarose gel electrophoresis. **B** Localization of in silico generated 10-mers in our data. Primary data analysis performed by Vedran Franke. **C** Validation of the 10-mer experiment by qRT-PCR

5.4.3. Using a transgenic reporter cell line to identify zip-codes in Ascl1-Ins

To increase the efficiency of reporter delivery into Ascl1-iNs and nascent RNA mimicry, the input library was cloned into the context of an artificial 3'UTR of a GFP reporter gene. This library pool was used to generate a transgenic cell line employing the piggyBac transposase system (Fig. 25a). Subcellular compartments of the reporter cell line were separated and RNA from these compartments was purified. All steps of cell line generation and reporter library extraction were controlled for complexity compliance. After adapter-specific purification of the reporters the libraries were PCR-amplified and the NextSeq workflow was followed using a read length of 150bp and one flowcell per compartment. Unfortunately, the extreme size and variety of sequences combined with multiple PCR steps, which were required due to low output of the assay, still resulted in inhomogeneity of the signal and therefore a low signal to noise ratio. The outcome was an extremely clustered sequence coverage (Fig. 25b). Therefore, peak calling algorithms could not be applied. Assuming the pool of sequences mimics a conventional differential expression experiment, using DeSeq2 for the analysis of the pools identified 69 transcripts as more than twofold enriched with a p-val below 0.05 in the neurite compartment thus being potential candidates to inherit localization signals.

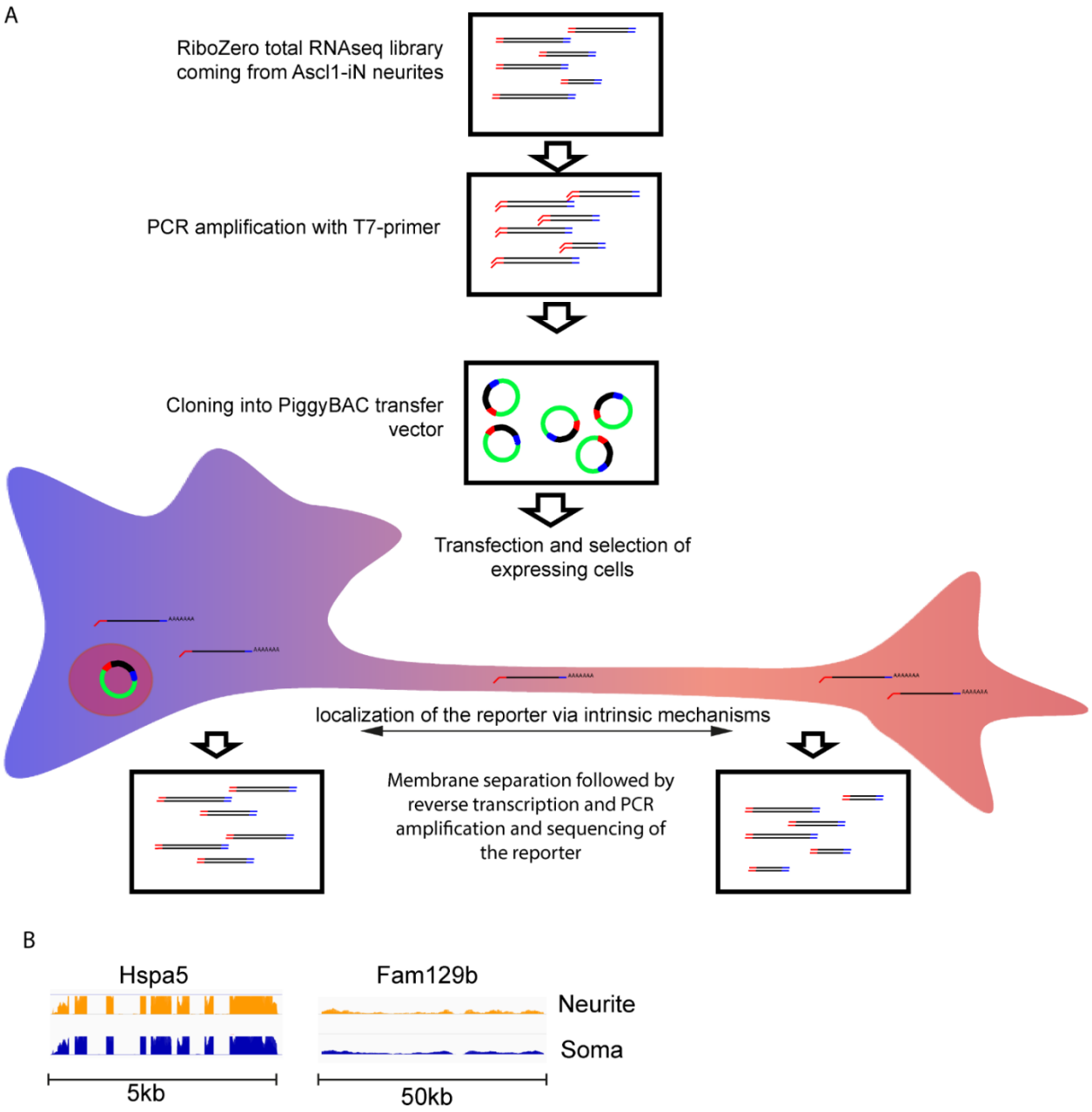


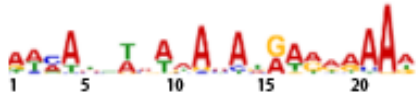
Fig. 25 Employing a polyclonal reporter cell line to identify zip codes in Ascl1-iNs.

A Experimental setup. The plasmid-based reporter library pool was generated via Gibson assembly. The cloning efficiency was validated by counting transformants throughout the cloning strategy. The final plasmid prep was performed using a pool of approximately 10 million transformants. 30 million Ascl1-mESC cells were transfected with 300µg of reporter plasmid. The selection was performed for four days on high dose to ensure purity of the polyclonal reporter cell line. 30 million GFP positive cells were plated on 30 membranes and total RNA was collected from subcellular compartments. The reporter library was extracted using RT-PCR and analyzed using conventional NGS techniques. **B** representative tracks of highly and lowly expressed genes. Primary data analysis performed by Esteban Peguero Sanchez.

To identify motifs that might drive the localization of these 69 mRNAs, MEME Suite database [315] and MAST [316] were used. Motive analysis of the enriched sequences revealed the

enrichment of an A-rich motive in the neurite pool when compared to the total pool or the soma enriched pool as a background model (Fig. 26a). The motif was found in 13 of the 69 neurite enriched mRNAs with an E-value below 0.1 and not at all in the soma pool (Fig. 26b). The motive was cloned into the 3'UTR context of the GFP reporter and used to generate a new reporter cell line. Analysis via qRT-PCR of the reporter construct did not show significant neurite enrichment (data not shown).

A



B

Gene name	Region	specific region	E-value
Ythdc1	inside_exon	inside_transcription_region	2.2e-6
Tnn7	inside_exon	3'UTR	3.7e-4
Gata4	inside_intron	5' UTR	4.9e-4
Fam63b	inside_exon	overlaps_3'UTR	9.2e-4
Ptp4a2	inside_exon	3'UTR	9.9e-4
Zfp942	inside_exon	3'UTR	3.7e-3
Cfap20	overlaps_exon_upstream	inside_transcription_region	8.4e-3
Zfp932	NA	NA	1.1e-2
Tceanc2	overlaps_exon_downstream	inside_transcription_region	1.7e-2
Fam46a	inside_exon	3'UTR	2.1e-2
Elovl7	inside_exon	3'UTR	3.4e-2
Insig1	inside_exon	3'UTR	4.4e-2
Gm9268	NA	NA	6.6e-2

C

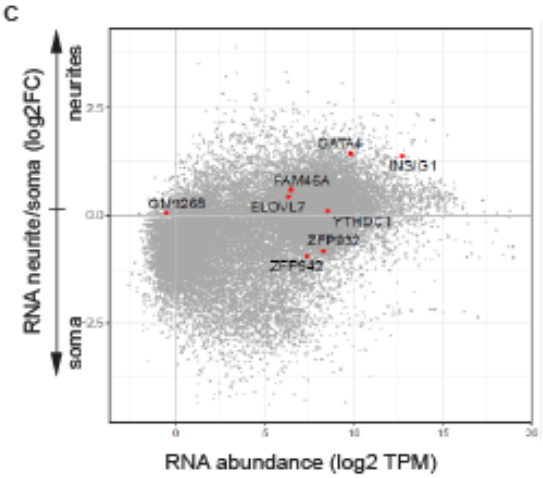


Fig. 26 Motive enrichment analysis of localized sequences.

A Identified motif from the neurite enriched reporter pool. Motifs were identified by MAST and MEME. Primary analysis was performed by Esteban Peguero Sanchez. **B** List of genes that inherit the found motif. **C** Mapping the genes in our initial spatial RNA-seq data. Eight out of 13 potential candidates inheriting the motive are also present in our initial dataset. Six of them show a positive neuritic enrichment, whereas two show a somatic enrichment in our initial RNA-seq data. There is no known RBP binding this motive.

5.4.4. The RNA element selection assay (RESA) provides the tools for precise mapping of cis-regulatory elements of RNA localization

The RNA element selection assay (RESA) was published by Yartseva et al. in 2017. Authors showed that it is suitable to identify regulatory RNA sequences inside a reporter pool during maternal-zygotic transition of a zebrafish egg in a qualitative and quantitative way. The limited size of the reporter pool is thereby crucial to gain the required tiling and precise mapping of regulatory sequences [284]. In theory this experiment can also work in identification of zip-codes in our test system. Since the delivery of a reporter to our cells was optimized using the PiggyBac transposon system, a RESA-suitable reporter library was cloned into the transfer vector.

Regulatory sequences of RNA localization (e.g. zip codes) were assumed to be in the 3'UTRs of localized transcripts. Therefore, the oligos for PCR amplification of the input library were

designed with the Primer3 [322], [323] software to span the 3'UTRs of 384 transcripts from the most neurite enriched mRNAs, in neurites of the initial spatial RNA-seq experiment and the 3'UTRs of proteins that showed more than twofold neurite enrichment even though the corresponding transcript was not enriched (see materials and methods - oligos). As an internal control and subsequent comparison to published data, the 3' UTRs of the top 48 enriched mRNAs from the Cajigas et al. [249] were also PCR-amplified and introduced to the reporter pool.

All PCR reactions that gave a product of expected size were pooled corresponding to the amount of amplicon and batch purified via AmpureXP bead purification (Fig. 27b). In total 364 fragments from our dataset and 41 fragments from the Cajigas et al. neuropil dataset could be amplified (see materials and methods - oligos). The resulting cDNA library was fragmented via sonication. Fragments were treated with the Klenow Large Fragment (NEB 0210) for 30min at 12°C and ligated sequencing adapters following the Illumina TruSeq stranded LT sequencing kit (steps A-tailing and adapter ligation). This fragment library mix was cloned into the 3'UTR context of a GFP reporter plasmid via Gibson assembly (Fig. 27a).

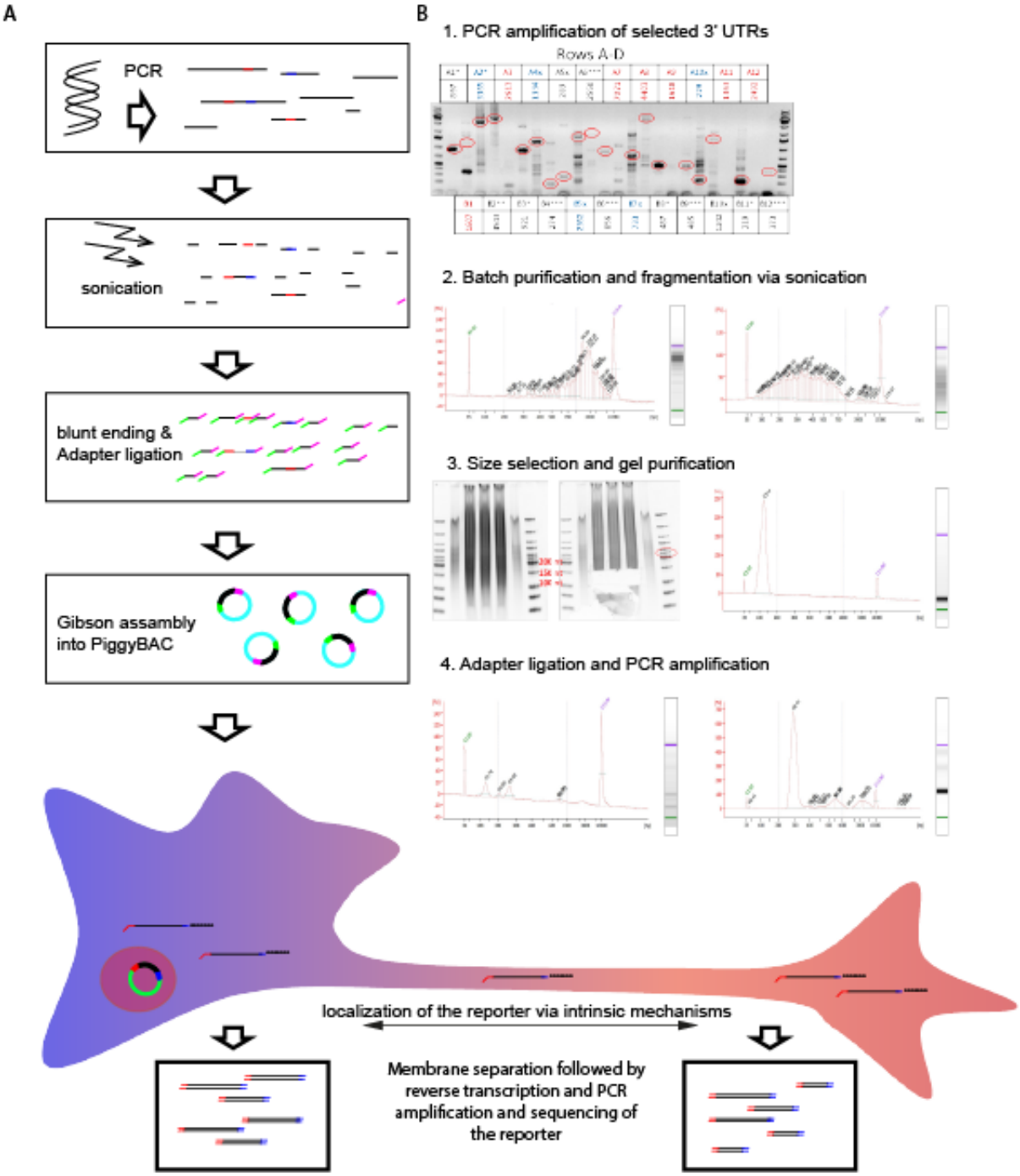


Fig. 27 Customizing the RESA protocol to identify zip-codes in neurons.

A Experimental setup. To ensure sufficient coverage of the reporter library it is important to monitor the number of fragments cloned into the plasmid-based RESA library, as well as the amount of incorporation events in the generated cell line. To reduce PCR bias, for each PCR step, the minimal amount of cycles was detected using qRT-PCR. C Oligo design for the generation of the RESA library. 405 candidate genes were amplified. Candidates were chosen from our identified neuritic enriched transcripts and from a validation dataset (Cajigas et al.). **B** RESA library generation. 1. Example of PCR reactions. PCRs were all performed in 96-well format using NEB Phusion high fidelity polymerase. The annealing gradient for each plate was determined by the highest and the lowest annealing temperature. The amplification time for each plate was chosen based on the longest transcript. Amplicons were pooled in comparable amounts based on agarose gel electrophoresis. Amplicons with many

unspecific products were cut from the gel, pooled as gel slices and batch purified using conventional gel purification techniques. All amplicons were pooled and purified using AmpureXP beads. 2.-4. All steps were monitored using a bioanalyzer with DNA1000 chip to ensure correct size of the fragments. To prevent PCR bias for all PCR steps a diagnostic qRT-PCR was performed in advance. 3. Size selection was achieved through gel purification excision and purification.

5.4.5. Generation of a plasmid-based RESA library

The plasmid library was used to generate a polyclonal transgenic reporter cell line. 30 million cells were transfected with 250µg GFP-reporter-transfer vector and 50µg of PiggyBac transposase expressing vector. Transfected cells were recovered for one day under normal growth conditions and were then selected for positive integration via hygromycin selection. To recover the cell numbers in the transgenic cell line, the cells were cultured under normal growth conditions for two more days prior to differentiation. To ensure high complexity of the library, the cell population was kept at a minimum of 30 million cells throughout the whole experiment. After separation of subcellular compartments and subsequent RNA extraction the reporter transcripts were isolated from endogenous RNAs via targeted RT and PCR amplification. These libraries were used to proceed with a standard sequencing protocol. Three replicates per compartment were sequenced. 276,249,538 reads for the neurite sample and 154,890,476 reads for the soma sample could be mapped uniquely to the genome. The overall distribution of reads showed a strong bias towards the 3' UTR, which is of course not surprising due to the selection of input fragments (Fig. 28b). Principle component analysis of the replicates shows a clear difference between the compartments (Fig. 28c). To identify peaks, e.g. sequences that are enriched or depleted in the neurite compartment, we used the MACS2 package [324], which identified 105,627 regions. Using DESeq2 [325] to obtain differentially localized sequences and filtering for those sequences with at least 60 reads in three replicates, considering only those sequences that we initially put in the library, revealed 139 regions being enriched in the neurite samples and 41 regions being enriched in the soma sample.

Mapping the sequences to the genome revealed a strong bias towards the 3' ends of the input sequences (Fig. 28d). This indicates a problem in the end repair and adapter ligation, during the initial library preparation.

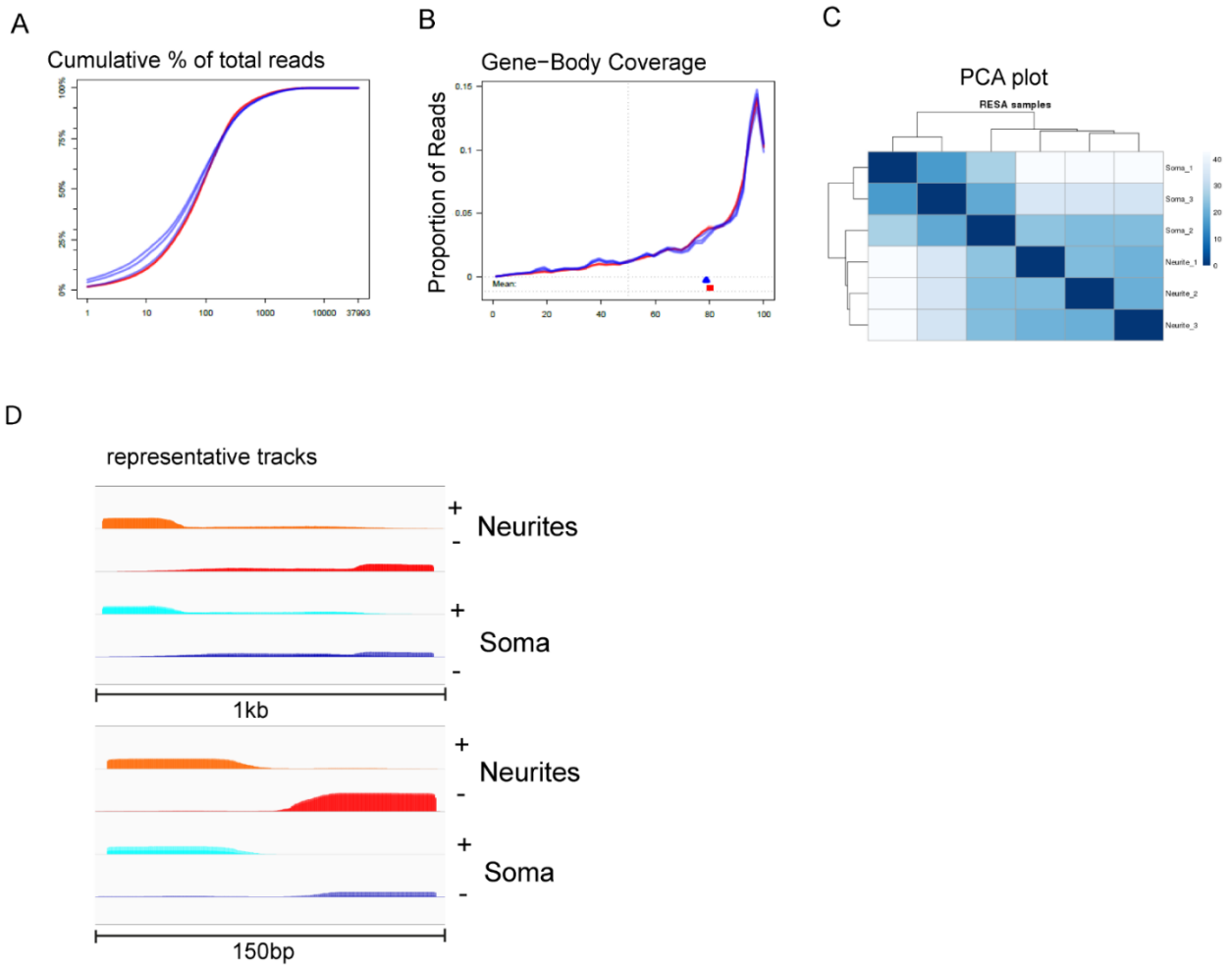


Fig. 28 Diagnostics of the plasmid-based RESA experiment.

A Cumulative distribution of reads over the genome. For all libraries about 75% of the reads map to 400 genes. Primary analysis was performed by Esteban Peguero Sanchez. **B** Cumulative distribution of reads over the gene body. The vast majority of reads maps to the 3' end of the transcripts. **C** Hierarchical clustering of the samples. There is a clear difference between neurite and soma samples but the sample "Soma 2" also overlaps with the neurite samples. **D** Representative tracks of two cloned fragments. A clear pileup of reads at the 3' ends of each construct is visible. The coverage between those "peaks" is very low.

6. Discussion

6.1. The neuronal model systems N1E-115 and Ascl1-iNs

Many approaches have been made recently towards identification of the dendritic/axonal transcriptome. Before 2016 most of these were either microarray-based approaches, yielding a maximum of 285 enriched mRNAs [241], [242], synaptosome analysis with all known impurities, or massive high-throughput in situ hybridization screens [239], identifying 68 mRNAs in the synaptic neuropil. Overlaps between those studies, however, yield a surprisingly small number of mRNAs discovered in more than one study. This might suggest that either identification is far from being saturated or the local transcriptome is extremely variable. Including studies from Taliaferro, Ainsley, Cajigas and Shigeoka enlarges the pool of potential synaptic mRNAs by far, although the overlap stays quite low (Fig. 29). Possible glia contaminations in samples from *in vivo* mouse brains might further bias these studies. It seems obvious that different cell types require different spatial transcriptomes due to the distinct molecular functions they must fulfil. In addition, Shigeoka et al. [109] showed that the local translatome changes significantly during embryogenesis and Ainsley et al. [151] showed that even in mature neurons the synaptic translatome is variable upon stimulation. Although both studies focus just on the ribosome bound mRNAs, these changes might as well be reflected in overall localization of the transcripts.

Using cell culture-based systems circumvents at least some of these variations. The used model systems here differ dramatically from each other. While N1E-115 cells are known to differentiate into active neuroblastoma cells [282], [285], [286], Ascl-mESC are known to build a homogenous population of post-mitotic Ascl-iNs [326]. Since no neuronal stimulatory supplement was used throughout this study it cannot be ruled out that the local transcriptome, translatome and proteome is as variable in our cell lines. It has been shown previously that cultured neuronal cells undergo spontaneous activation [327] and, therefore, our cell population might consist of a mixture of activity states. This might be a disadvantage for the investigation of specific mechanisms of neural plasticity, though it does not intervene with identification of the stable “steady-state” of the spatial molecular environment.

The local transcriptome and translatome of neuronal cells also seem extremely variable and highly cell type specific in the data of this study. Comparing the spatial transcriptome data of N1E-115 and Ascl1-iNs as well as other published datasets of cortical neurons, CAD neurons, N2A neurons to an *in vivo* mixture of CA1 neurons, there is a marginal overlap (Fig. 29). It is clear that different cell types with different cellular functions require specific subcellular environments.

The missing correlation between RNA and protein localization in N1E-115 might be an outcome of technical difficulties during the experiment.

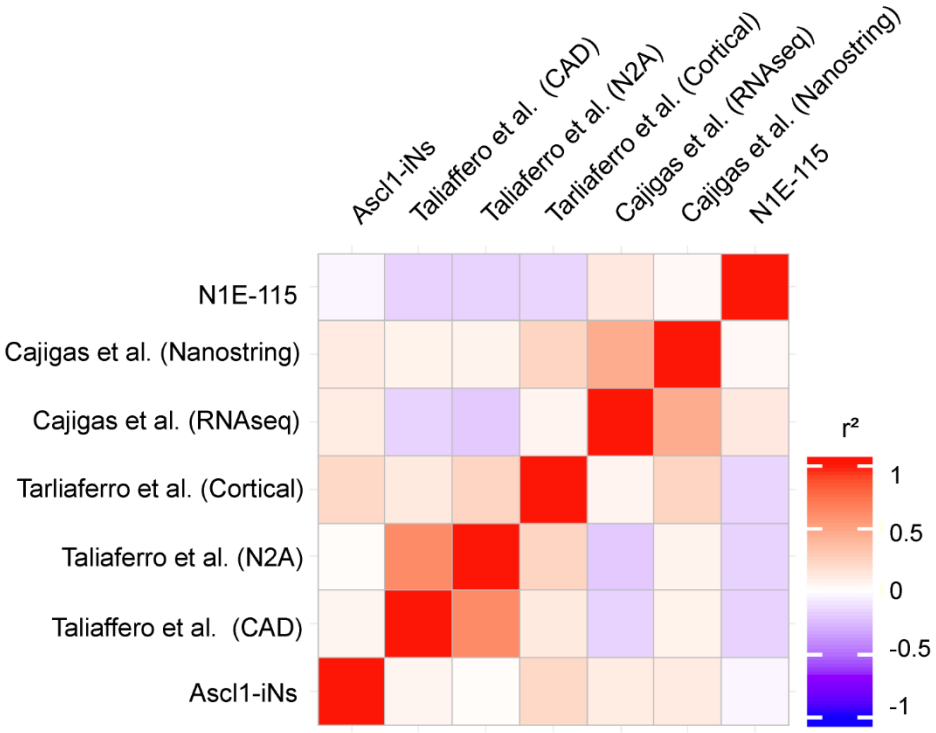


Fig. 29 N1E-115 and Ascl1-iNs data in correlation to other publications. The correlation of diverse spatial RNA-seq experiments is very weak. The usage of different cell types, different developmental stages and different sequencing conditions might be the reason for this phenomenon.

6.2. The role of RNA localization and local translation in establishing local protein pools in neurons

The existence of spatial protein production in neurites and the role of mRNA localization has been a hotly debated topic in neuronal cell biology. It has been shown that local translation is crucial for neuronal plasticity [72], [152], [328], neuronal survival [329], synaptic differentiation [330], as well as for differentiation [331] cell migration [332], the cell cycle [333] and cellular polarity [39], [45], [331], [334]. Despite the implication of spatial translation in specific pathways, the global distribution has never been shown. An *in vitro* subcellular neuronal fractionation system was employed for this study, in combination with mass spectrometry and RNA-seq to quantify the transcriptome-wide patterns of mRNA localization, the impact of mRNA localization on local translation in Ascl1-iNs. Additionally, a translatome dataset [307] of the same cell line was used to highlight the impact of spatial translation. Using a minimal cutoff of twofold enrichment in the neurite datasets the analysis revealed that neurite-targeted mRNAs encode for approximately 63% of the spatial proteome (protein log2FC neurite/soma > 1; 148 out of 234 proteins; RNA log2FC neurite/soma > 1, P-values < 0.05). Using RiboSeq, pSILAC and QuaNCAT it can be confirmed that most of the localized

mRNAs are locally translated. Consistently with neurite localization, GO term enrichment analysis showed that these genes are associated with molecular functions “actin cytoskeleton”, “calcium ion binding”, “extracellular matrix,” and “growth factor binding”. Furthermore, a substantial group of the localized proteins are encoded by mRNAs, which are moderately enriched in neurites (79 out of 234 proteins; 0 < RNA-seq log2FC < 1). These proteins may represent an intermediate group localized via multiple mechanisms, involving both mRNA and protein transport and differential translation efficiencies. Calculating these translation efficiencies (diff_trans log2FC= RiboSeq log2FC / RNA-seq log2FC) reveals that more than 85% (197 out of 234) of the local proteome can be explained by local protein synthesis, leaving a very small part of the local proteome that cannot be explained by mRNA localization or neurite localized translation (4 out of 234 proteins; diff_trans log2FC and RNA-seq log2FC < 0; Fig.30). Mechanisms of protein transport may underlie those cases as well as temporal regulated translational regulation. Due to technical limitations those mechanisms could not be analyzed in this study.

Local protein synthesis has been implicated in several neuronal diseases, neurodevelopmental disorders, and psychiatric and degenerative diseases [148]. In fact, it was found that approximately 40% of the localized genes are implicated in neuronal diseases such as the DiGeorge syndrome, autism spectrum related disorders, fragile X syndrome, Alzheimer’s disease and Parkinson’s disease. Localization of these mRNAs suggests that dysregulation of either localization or local translation of these mRNAs may give rise to some of the mentioned disorders.

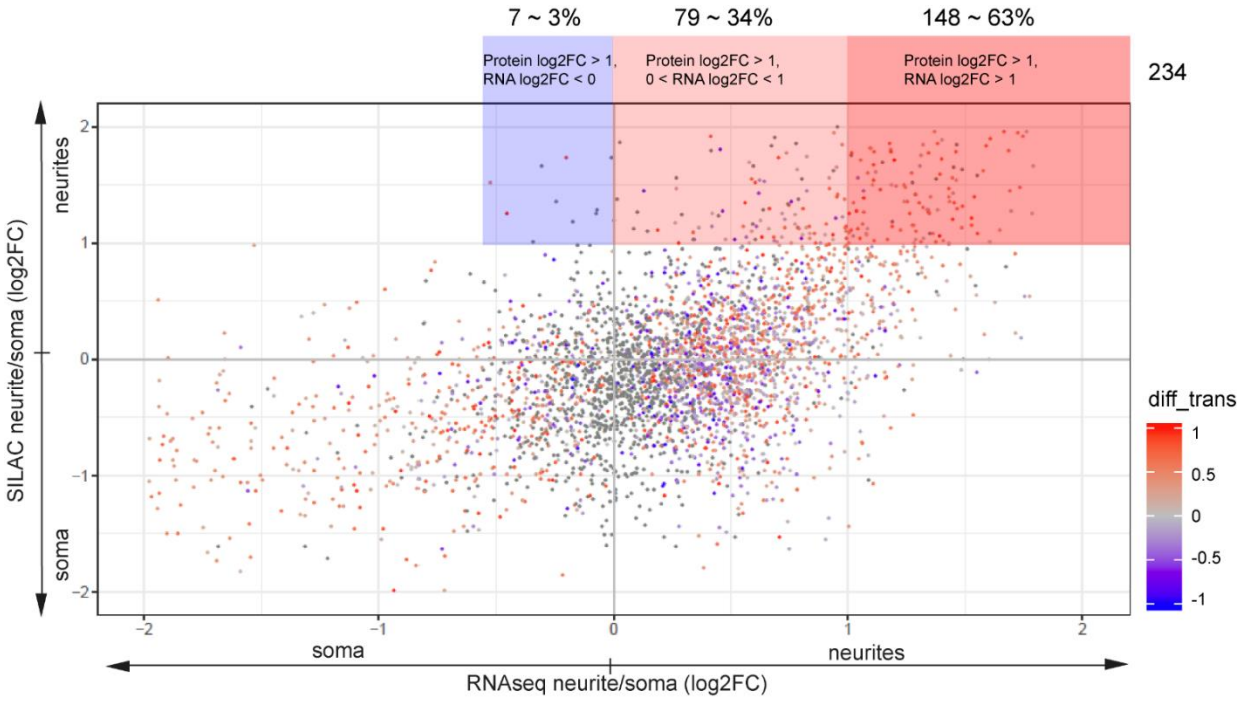


Fig. 30 Spatial translation is the main regulator of the local proteome. 63% of the neuritic enriched proteins are also enriched as mRNAs (SILAC REV loc2FC > 1; RNA-seq log2FC > 1 & p-val < 0.05). 34% of the spatial proteome is at least slightly enriched as mRNAs (SILAC REV loc2FC > 1; 0 < RNA-seq log2FC < 1 & p-val < 0.05). 3% of the local proteome correspond to somatic enriched mRNAs (SILAC REV loc2FC > 1; RNA-seq log2FC < - 1 & p-val < 0.05). Three of those proteins do have a positive differential translation efficiency, indicating a more efficient spatial translation in the neurite.

Correct subcellular localization of proteins is crucial for neuronal functions. It can be achieved (1) by transporting proteins with molecular motors as parts of RNPs or vesicular organelles, (2) through mRNA localization and local translation, or (3) via preferential local translation of equally distributed mRNAs, i.e., due to localization-dependent translational regulation. Specific examples for each mechanism have been described in the literature, but it is unclear to what extent each mechanism contributes to the overall protein distribution asymmetry. One reason for this is that most genome-wide studies focused on one particular layer of gene expression - transcriptome, proteome, or translatome, or a single cellular compartment (e.g., axon profiling without comparison with the somatic compartment). For example, Taliaferro et al. [81] applied RNA-seq to neurites and the soma of neuronal cell lines and mouse cortical neurons to study the differential localization of RNA isoforms. Shigeoka et al. [109] used the Ribotag approach to identify ribosome-bound mRNAs in mouse retinal axons.

Due to differential sensitivity of high throughput mass spectrometry and sequencing methods, the proteome and translatome data vary in the number of detected transcripts. Because of this differential detection threshold, there is a pool of mRNAs that is enriched in the neurite

compartment but not translated, or is translated but is not identified on the protein base by the experiments in this study (low abundance, no specific peptides). This pool also contains RNAs that stay translationally repressed and might be released upon a specific stimulus. As it was shown by Ainsley et al., a big part of the neurite ribosome associated transcriptome changes upon intensive neuronal stimulation [151]. To understand the function of axonal translation and mRNA localization it is important to carry out comprehensive and unbiased global studies in various homogenous, synchronized neuronal populations.

Of course, this raises the question of how the local translatome is controlled by upstream inter- and intra-cellular signaling. It was postulated that mTorc1 regulates axonal translation [335]. mTorc1 is important for axon organization [336] and also peaks in actively wiring axons [109]. Its targets show a steep increase at P0.5 and many of them are also found in the neurite dataset. In contrast to this, the activity of Fmrp was predicted to peak later in development since target mRNAs show a coordinate decrease in ribosome association in mature axons. The translational break mediated by Fmrp is also utilized more in the mature CNS [337], [338].

However, the translational regulator APC, which is known to regulate microtubule assembly, axonal growth and plays a critical role in local translation in various systems of axis formation and neurogenesis, seems to be most important in immature axons at stage E17.5 [109].

A recent study on mRNA and protein localization in MDA-MB231 breast cancer cells did not find any correlation between mRNA and protein localization [295]. This suggests that mRNA localization and spatial translation is more critical in highly polarized cells with long protrusions such as neurons. It was also shown that cancerogenous cells can suffer from significant abrogation in mechanisms of mRNA transport and spatial translation. Cancer cell lines are therefore an imprecise model for the exploration of *in vivo* mechanism of spatial localization. The lack of spatial translatome data in these migrating cells makes a comparison to our system impossible. It might be that outgrowth of migrating cells holds completely different localization machineries than steady state differentiated neurons. The establishment of the spatial translational machinery including ribosomes, translation factors, regulatory elements and posttranslational modifiers might not be energetically rewarding for a cell that is continuously changing its shape, although it is advantageous for the long distances and established immotile compartments of terminally differentiated neurons. This hypothesis is consistent with the observation of no correlation between protein and RNA localization in N1E-115 cells.

Emerging data indicates that local protein synthesis is crucial for synaptic development and plasticity [117], [339], [340], although protein production in the cell bodies is undoubtedly

important. The data suggests local protein synthesis is the major player of building the local proteome of cultured iNeurons. It is of undoubted importance to gain comparable datasets from other model systems including *ex vivo* primary neurons and *in vivo* experiments to validate these results in the mammal brain.

The revealed translatome includes several neurotransmitters, receptors, adhesion molecules, signaling molecules and regulators of protein synthesis, but also several transcription factors and other proteins associated with nuclear functions. The phenomenon of axonal translation and retrograde trafficking was already examined for diverse transcription factors like Creb [259], Smad1, 5 and 8 [341] and Stat3 [342]. As already suggested by others there are at least two explanations for this phenomenon: (1) though previously associated with nuclear functions, many proteins might have additional functions; and (2) the distal translation and retrograde trafficking of transcription factors and DNA modifiers implies one additional layer of intracellular signaling. As post-translational modifications are often achieved co-translationally they rely heavily on the molecular environment. These modifications further alter the function of a nuclear protein as it is trafficked into the nucleus, depending on its NLS. Through this mechanism intercellular signaling is translated rapidly into transcriptional regulation.

Despite genes associated with synaptic functions, we can find several axon-related proteins with clear spatial localization in our dataset. Although we cannot distinguish in our data between axonal and dendritic localization and translation, this is another hint towards axonal protein synthesis in mature neurons. This phenomenon was long doubted and could only be measured in developing axons [343], [344]. Other recent publications have also discovered a wide range of spatially translated proteins in mature axons [109], [246]. These studies suggest that axonal mRNA translation continues in the mature brain and may regulate presynaptic functions changing during synaptic excitation and over the lifetime of the neuron.

The presence of mRNAs for diverse membrane proteins and their translation including several receptors and ion channels is not surprising. Many of those channels are known to be expressed in gradients from the soma towards the dendrites supporting excitatory potential of the dendrites and local control of signaling [345], [346]. The local synthesis of voltage gated channels and its regulation via neuronal excitation implies another layer of neuronal signaling. It was shown by Raab-Graham et al [347] that Kv1.1 translation is suppressed by synaptic excitation in pyramidal neurons, resulting in enhanced excitability [347]. Similar feedback loops are at least possible, if not probable. Multiple omics datasets from ground state and excited neurons would be of undoubtful interest to solve this puzzle. The fact that membrane

associated proteins are observed as being translated in the distal cellular compartments naturally requires the presence of the co- and posttranslational processing machinery of these proteins. Although it was known before that also distal parts of neurons contain components of the ER and Golgi [348]–[351], it remained a matter of debate whether the whole secretory pathway, the glycosylation pathway, and also key enzymes that influence ER export are present. Components of all of those machineries were identified in the neurite dataset of this study. Although most of them do not seem to be enriched in the neurite compartment, they are clearly present in the neurite. The glycosylation of membrane proteins is especially crucial for their correct folding and function.

This study provides first global insights into the molecular architecture of outgrowing neurites on the proteome, transcriptome and translatome level. It is a valuable resource for future analysis of neuronal functions. Of course, mRNA localization and local translation is not only crucial for neuronal functions; many mechanisms of localization are also pivotal for cell migration, the establishment of cellular symmetry, cell cycle functions, cellular differentiation, especially in the context of inequal cell divisions. Also intercellular communications, including cell-cell contacts, cell-ECM contacts and mechanisms of RNA localization as well as spatial translation will not be exclusive to neurons.

6.3. Spatial translational regulation in outgrowing neurites

The directed outgrowth of axons is triggered by various extracellular cues inducing elongation or collapse of the growth cone. It was shown that inhibition of translation prevents sensitivity to most of these external signal [335], [352]. While attractive factors like netrin-1 or nerve growth factor (Ngf) stimulate spatial protein synthesis of cytoskeleton associated genes in the axon growth cone [353]–[355], repulsive factors like Sema3A or Slit2 stimulate the spatial disassembly of the cytoskeleton [161], [356]. The external signal gradient induces filopodia and lamellipodia formation and withdrawal and, thereby, axon architecture and navigation [353], [354]. Additionally, it was shown that various components of intracellular signaling pathways, guidance receptors and cell adhesion molecules are translated in neurites under vigorous control [109], [357].

After reaching their terminal target through the molecular guidance gradient, axons initiate branching to connect via synapses with their post-synaptic partners [358]. It was shown that branching points are points of protein synthesis *in vitro* [336], [359], [360] and destinations of mRNP transport [361]. Inhibition of translation inhibits both initiation and stabilization of new branching points [361], suggesting that regulation of spatial protein production is crucial for this mechanism. Finally synaptogenesis requires local translation in the presynaptic

compartment [362]. Spatial synthesis of Snap25 and β-catenin is crucial for presynaptic vesicle release and, therefore, neuronal functionality [362]–[364]. It was shown that GABA release in established synapses and, therefore, long-term synaptic plasticity underlies controlled local protein synthesis. The question of how thousands of different mRNAs are spatiotemporally controlled in translation is of high impact for the field of neuroscience, since it bears answers to fundamental questions of neuronal functions. Many axon guidance cues and neurotrophins are known to stimulate kinases that further induce phosphorylation of specific RBPs. Through this mechanism target mRNAs can be released, degraded, modified or bound and translational patterns can be changed [365]–[368]. A prominent example of this is the mTORC1 pathway that selectively promotes cap-dependent translation of target mRNAs. The fine-tuning of multiple pathways like these could bear the potential of defining spatiotemporal control of protein synthesis.

As RBPs are crucial for RNA metabolism and translational regulation in every direction, I focused also on neurite enriched RBPs. It is obvious that neurite enriched RBPs bear the potential to mediate mRNA localization as well as regulate spatial translation and spatial mRNA turnover.

The Elav like protein 2 (Elavl2), for example, was identified as a neurite enriched RBP. It was shown that this protein is implicated in mRNA localization in the *Xenopus* egg cell during oogenesis [369]. Additionally, it was shown to be important during neuronal differentiation and its dysregulation is linked to severe neuronal diseases like autism [370]. Also, the nuclear export factor 7 (Nxf7) which was shown to be additionally involved in RNP granule formation in neurons and is therefore linked to neuronal RNA localization [371], was found enriched in the data.

RBPs with known roles in mRNA decay and translational regulation such as Mov10 and Lsm14b were also found, which could regulate the levels of neurite targeted transcripts and their translation efficiencies. Several RBPs which are known to play a role in protein transport like Mvp or Rrbp1 are, of course, interesting candidates for simultaneous protein and RNA localization.

Peg10 is a neurite enriched RBP that is known to be involved in cell growth promotion and interacting with the tissue growth factor beta (Tgf-β) [372].

Neurite enriched ribosomal proteins like Rpl26, Rpl27 and Rpl31 support the hypothesis of distinct spatial ribosomal subpopulations that could enable the cell to regulate spatial translation [373]. In particular, the presence of Urb2, a ribosome biogenesis factor, and Tfb1m, a protein implicated in rRNA modification, supports this idea. Spatial pools of “specialized”

ribosomes could react differentially to ex- or internal stimuli and could drive differential spatial translation.

Other enriched RBPs are known to be involved in general RNA metabolism and might therefore be important for spatial RNA turnover rated regulating spatial translation indirectly. Examples for this group are RNaseH2c, Rexo2 or Hrsp12.

Another interesting group of neurite-enriched RBPs are transcription factors and DNA regulators. We found Ttf2, Pura and Gtf2f enriched in the neurite fraction. Both Ttf2 and Pura are gene specific transcription factors known to be involved in neurological disorders [374], [375], whereas the general transcription factor, 2 subunit 2f (Gtf2f), is a general transcription factor promoting transcription initiation. The presence of transcription factors in distal parts of the cell like the neuronal growth cone was already observed by Cox et al. in 2008 [259]. It was postulated that locally translated transcription factors can be modified depending on the molecular environment of their synthesis and that these modifications might resemble a feedback loop altering their functions when they are retrograde trafficked to the nucleus.

The specific functions of these RBPs are under investigation, using knockout studies in combination with genome-wide target identification, using RNA immunoprecipitation (RNA-IP), or crosslinking and immunoprecipitation (CLIP) assays to provide additional insight into the mechanisms by which those RBPs establish cell polarity and neuronal functions.

Locally enriched miRNAs can also be implicated in differential spatial translation as well as differential mRNA localization. The miRNAome of neurons and neurites is highly cell type and most probably also development specific [376]. Intracellular miRNA gradients of course bear the potential of altering spatial protein synthesis rates, as does alternative 3' UTR – localization. The observed trend of longer 3' UTRs in distal neuritic compartments by Taliaferro et al. [81] could hint towards spatial regulation via 3'UTR-binding trans factors like RBPs and miRNAs. There are examples known for miRNAs controlling axonal synthesis of target mRNAs like miR-338 [199], [279] and also spatially responding to signaling as miR-182 in growth cones [193]. Localization and local processing of miR precursors was demonstrated in dendrites [189] and is still under debate in axons.

In total 191 miRNAs were identified as preferentially localized to neurites in Ascl1-iNs. Some of them are known to exhibit neuronal functions and to be localized to dendrites or axons like members of the let-7 family [311], [312] and miR-483 [308]. The overlap with a published dataset of axonally enriched miRNAs of sympathetic neurons is rather poor. Three out of 17 axonally localized miRNAs found by Natera-Naranjo et al. [199] could be identified in our

neurite enriched data. Of course, this might be due to the nature of miRNAs that are highly cell type specific.

Using in silico miRNA-target predictions based on experimental data we could not observe any global effect on either localization or translation of miRNA targets in our assay, though it should be mentioned that these effects are often rather small and difficult to observe in global data. A global miRNA-target interaction database for Ascl1-iNs is crucial to identify direct and indirect effects of localization and translational regulation in a dataset like this one. In 2014 Sasaki et al. published a dataset of axon and growth cone enriched miRNAs [200]. This dataset was generated using massive qRT-PCR analysis of subcellular fractions from *in vitro* cultured hippocampal neurons. It is postulated that despite their role as fine tuners of gene expression, miRNAs are key regulators of neuronal function and development [377]. In total six miRNAs were found enriched in their axonal fraction. Obviously, that dataset is not complete due to the nature of the experiment and the overlap with the data of this study is therefore very poor.

In silico examination of miR influence showed that miR-1 targets decrease during axon maturation in RGC development [109]. It might be that most miRNAs are (1) activity dependent or (2) active during maturation and, thereby, the effects of miRNAs in the test system of this study could not be observed.

Posttranscriptional modifications are another wide field of spatiotemporal translational regulation. Axonal *Gap43* mRNA is modified, for example, with n6-methyladenosin (m6A) [378] and, therefore, a substrate for the demethylase enzyme fat mass and obesity-associated protein (Fto). Depending on its methylation state, local translation is either promoted or repressed [378]. Unfortunately, no global analysis of spatial mRNA modifications has been published yet. The direct coupling of extracellular signaling to ribosome dissociation and release for example via Netrin-1 [379] is another attractive potential mechanism for spatial, yet unspecific, translational regulation. Since some mRNAs coding for ribosomal proteins are enriched in the neuritic fraction, it is possible that local translation of ribosomal proteins can repair damaged ribosomes or alter composition of local ribosome populations [109], [366].

In contrast to other studies, we could not detect extensive isoform diversity using MISO [111] in our datasets. This might be due to the type of sequencing (small read length) and protocol of library preparation (not 3' end specific). Further studies including 3' end sequencing in subcellular compartments of Ascl1-iNs will be of great interest to identify the isoform specific localization and local translation. Also, MISO does not *de novo* detect exon junctions, but only uses predefined annotations. It might be that localized isoforms are not annotated and are therefore excluded from the analysis.

6.4. Spatial circular RNAs bear the potential for unknown functions in subcellular compartments of neurons

The functions of circular RNAs are, except for some specific examples, not known, althoughtheir abundance, tissue specificity, transcriptional regulation and conservation clearly points towards functional importance [380], [381]. They seem to be especially enriched in neuronal tissues [300].

In total 376 circular RNA isoforms could be identified in Ascl-iNs, of which 36 showed an enrichment in neurites (circular isoform in neurites / circular isoform in soma > 2) and 61 showed a circ specific enrichment (circular isoform neurites / linear isoform neurites > 2). Additionally, seven genes showed an enrichment of the circular isoform, while the linear isoform is depleted from the neurites (circular isoform neurites/ circular isoform soma > 2 & linear isoform neurites/ linear isoform soma < 2). It is known that neuronal cells are enriched for circular spliced isoforms and that circular spliced isoforms are found in synaptosomes and axons of terminally differentiated neurons [381]. Here it is shown for the first time that some circular isoforms are enriched in the neurons of Ascl1-iNs. This enrichment seems to rely on their circular nature since the linear transcript of many of these circRNAs is not enriched. It might indicate specific functions of the circular isoforms in the subcellular compartment of these cells. There are few circular RNAs for which the function is known. Despite others, one reason for this is the complicated validation of circular-specific functions. *In vivo* most of these circular isoforms are very low abundant,although most hypothesis are based on batch sequencing experiments, therefore average the found molecules over a large number of cells and of course also over the whole cell body. Spatial enrichment of circular isoforms bears the potential to alter these hypotheses, since the relative abundance of said circular isoforms in the subcellular compartment might exceed the homeopathic concentrations calculated by canonical bulk RNA-seq experiments. Interestingly, the top enriched genes, for which the circular isoform is specifically enriched, are associated with polarization-related terms. *Ptpn14* [382] and *Mob3b* [383] are associated with receptor kinase activities and involved in signal transduction. *Lama1* [384] and *Fndc3a* [385] are associated with cell-cell adhesions, while *Zwilch* [386], *Spag5* [387] and *Ephb2* [305] are associated with cytoskeletal function, protein transport and axon guidance, respectively. However, the circular isoform of a gene canonically involves just the first 2-4 exons and translation of these circular isoforms was not yet validated in the system used in this study.

6.5. Identification of cis-regulatory elements and trans acting factors of RNA localization in Ascl1 iNs

The importance of RNA localization for cellular life and especially highly polarized cells such as neurons is indisputable. Thus, the dissection of mechanisms of RNA localization is as important as it is challenging. Here the combination of a multi reporter based, high throughput method for the *de novo* identification of zip codes in neurons is proposed. Knowledge about how and why the message is delivered to a specific compartment bears the potential to design target-specific drugs in the form of artificial messengers that work on a subcellular level. For molecular biology the knowledge about what goes where is also of crucial importance, since it enables the user to make subcellular compartment specific reporters for other mechanisms of polynucleotides and proteins.

The global analysis of mRNA localization is, however, a very broad field and many attempts have been made to identify players of mRNA localization. The presented approach focuses now for the first time not on the trans acting factors that might be involved, but the cis-acting sequence that might inherit the potential to be localized. Most known RBPs that do contribute to mRNA localization also fulfil other molecular functions, therefore pinpointing the effect only on mRNA localization from RIP or CLIP data is problematic. Furthermore, follow-up knockout or knockdown studies of specific RBPs don't provide clear answers to the question of localization, since the secondary effects of these studies might hide localization mechanisms. Due to the sophisticated mechanisms that interplay in the orchestra of mRNA localization (e.g. splicing, degradation, stabilization, direct transport) the term "zip-code" appears to be system specific. An RBP that is involved in overall transcript degradation in cell type A might also be involved in spatial degradation and therefore localization of the same transcript in cell type B.

6.5.1. Problems and solutions of the RESA-approach

The delivery of the reporter library turned out to be crucial in this system. Yartseva et al. [284] used microinjections into the *Danio rerio* embryo, which is of course not suitable for this model system. RNA transfections might be a valid method, although the transfection efficiency of post-mitotic neurons is not very high and, therefore, the amounts of recovered reporter mix required more PCR steps prior to the sequencing, which resulted in less complexity of the output and thus a reduced tiling.

The second problem of this approach was the extreme size of the input library (whole transcriptome), which resulted in a crude tiling. Since the fine and overlapping tiling of the reporter library is of crucial importance for bioinformatical analysis, the complexity of the input library must be reduced (as it is stated in the original RESA paper). The block-like over-

amplification of few sequences, which prohibited canonical peak-finding algorithms to work on this data could be avoided by reducing the library size as well.

Reducing the reporter library to selected candidates as it was done in the original paper is crucial for the generation of a reasonable signal to noise ratio. Unfortunately in this attempt, either the blunt ending reaction or the A-tailing reaction failed, resulting in a strong bias in adapter ligation towards the 3' ends of the constructs. The inequal coverage of the library prevented identification of RNA elements that might drive localization.

7. References

[1] B. Alberts *et al.*, *Alberts*. 2009.

[2] J. M. Slack and D. Tannahill, "Mechanism of anteroposterior axis specification in vertebrates. Lessons from the amphibians.," *Development*, vol. 114, no. 2, pp. 285–302, 1992.

[3] J. D. Halley, F. R. Burden, and D. A. Winkler, "Stem cell decision making and critical-like exploratory networks," *Stem Cell Res.*, vol. 2, no. 3, pp. 165–177, 2009.

[4] J. D. Halley, K. Smith-Miles, D. A. Winkler, T. Kalkan, S. Huang, and A. Smith, "Self-organizing circuitry and emergent computation in mouse embryonic stem cells," *Stem Cell Res.*, vol. 8, no. 2, pp. 324–333, 2012.

[5] K. Ebnet, D. Kummer, T. Steinbacher, A. Singh, M. Nakayama, and M. Matis, "Regulation of cell polarity by cell adhesion receptors," *Semin. Cell Dev. Biol.*, 2017.

[6] W. Zhang *et al.*, "Nucleosome positioning changes during human embryonic stem cell differentiation," *Epigenetics*, vol. 11, no. 6, pp. 426–437, 2016.

[7] B. Alberts *et al.*, *Molecular Biology of the Cell 6e*, vol. 6, no. 6. 2014.

[8] J. A. Knoblich, "Mechanisms of Asymmetric Stem Cell Division," *Cell*, vol. 132, no. 4. pp. 583–597, 2008.

[9] J. Nance, "Getting to know your neighbor: Cell polarization in early embryos," *J. Cell Biol.*, vol. 206, no. 7, pp. 823–832, 2014.

[10] D. S. Johnston, "Establishing and transducing cell polarity: common themes and variations," *Curr. Opin. Cell Biol.*, vol. 51, pp. 33–41, 2018.

[11] H. Harashima, N. Dissmeyer, and A. Schnittger, "Cell cycle control across the eukaryotic kingdom," *Trends Cell Biol.*, vol. 23, no. 7, pp. 345–356, 2013.

[12] a Udvardy, "The role of controlled proteolysis in cell-cycle regulation.," *Eur. J. Biochem.*, vol. 240, no. 2, pp. 307–13, 1996.

[13] M. Malumbres, "Cyclin-dependent kinases," *Genome Biol.*, vol. 15, no. 6, pp. 1–10, 2014.

[14] M. Donzelli and G. F. Draetta, "Regulating mammalian checkpoints through Cdc25 inactivation," *EMBO Rep.*, vol. 4, no. 7, pp. 671–677, 2003.

[15] S. J. Rahi, K. Pecani, A. Ondracka, C. Oikonomou, and F. R. Cross, “The CDK-APC/C Oscillator Predominantly Entrains Periodic Cell-Cycle Transcription,” *Cell*, vol. 165, no. 2, pp. 475–487, 2016.

[16] M. Cavey and T. Lecuit, “Molecular Bases of Cell–Cell Junctions Stability and Dynamics,” *Cold Spring Harb. Lab. Press*, vol. 10, no. 1101, pp. 1–18, 2009.

[17] S. Getsios, D. P. Kelsell, and A. Forge, “Junctions in human health and inherited disease,” *Cell Tissue Res.*, vol. 360, no. 3, pp. 435–438, 2015.

[18] D. E. Ferrier and P. W. Holland, “Ancient origin of the Hox gene cluster.,” *Nat. Rev. Genet.*, vol. 2, no. 1, pp. 33–38, 2001.

[19] L. L. McGrew, S. Hoppler, and R. T. Moon, “Wnt and FGF pathways cooperatively pattern anteroposterior neural ectoderm in Xenopus,” *Mech. Dev.*, vol. 69, no. 1–2, pp. 105–114, 1997.

[20] W. H. Palmer and W. M. Deng, “Ligand-Independent Mechanisms of Notch Activity,” *Trends Cell Biol.*, vol. 25, no. 11, pp. 697–707, 2015.

[21] K. Takahashi and S. Yamanaka, “Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors,” *Cell*, vol. 126, no. 4, pp. 663–676, 2006.

[22] X.-L. Guo and J.-S. Chen, “Research on induced pluripotent stem cells and the application in ocular tissues,” *Int J Ophthalmol*, vol. 8, no. 4, pp. 818–825, 2015.

[23] V. Kashyap *et al.*, “Regulation of Stem Cell Pluripotency and Differentiation Involves a Mutual Regulatory Circuit of the Nanog, OCT4, and SOX2 Pluripotency Transcription Factors With Polycomb Repressive Complexes and Stem Cell microRNAs,” *Stem Cells Dev.*, vol. 18, no. 7, pp. 1093–1108, 2009.

[24] L. R. Smith, S. Cho, and D. E. Discher, “Stem Cell Differentiation is Regulated by Extracellular Matrix Mechanics,” *Physiology*, vol. 33, no. 1, pp. 16–25, 2018.

[25] M. Yoshida, E. Muneyuki, and T. Hisabori, “ATP synthase - A marvellous rotary engine of the cell,” *Nat. Rev. Mol. Cell Biol.*, vol. 2, no. 9, pp. 669–677, 2001.

[26] R. S. Cohen, “The role of membranes and membrane trafficking in RNA localization,” *Biol. Cell*, vol. 97, no. 1, pp. 5–18, 2005.

[27] S. Backes and J. M. Herrmann, “Protein Translocation into the Intermembrane Space and Matrix of Mitochondria: Mechanisms and Driving Forces,” *Front. Mol. Biosci.*, vol. 4, no. December, pp. 1–11, 2017.

[28] S. Takamori *et al.*, “Molecular Anatomy of a Trafficking Organelle,” *Cell*, vol. 127, no. 4, pp. 831–846, 2006.

[29] T. Harayama and H. Riezman, “Understanding the diversity of membrane lipid composition,” *Nat. Rev. Mol. Cell Biol.*, 2018.

[30] D. A. Fletcher and R. D. Mullins, “Cell mechanics and the cytoskeleton,” *Nature*, vol. 463, no. 7280, pp. 485–492, 2010.

[31] N. Inagaki and H. Katsuno, “Actin Waves: Origin of Cell Polarization and Migration?,”

Trends in Cell Biology, vol. 27, no. 7. pp. 515–526, 2017.

[32] T. D. Pollard and J. A. Cooper, "Actin, a central player in cell shape and movement," *Science*, vol. 326, no. 5957. pp. 1208–1212, 2009.

[33] E. Vollmeister, K. Schipper, and M. Feldbrügge, "Microtubule-dependent mRNA transport in the model microorganism Ustilago maydis," *RNA Biol.*, vol. 9, no. 3, pp. 261–268, 2012.

[34] M. L. De Heredia and R. P. Jansen, "mRNA localization and the cytoskeleton," *Curr. Opin. Cell Biol.*, vol. 16, no. 1, pp. 80–85, 2004.

[35] D. Sept, "Microtubule Polymerization: One Step at a Time," *Curr. Biol.*, vol. 17, no. 17, pp. 764–766, 2007.

[36] H. Herrmann, H. Bär, L. Kreplak, S. V. Strelkov, and U. Aebi, "Intermediate filaments: From cell architecture to nanomechanics," *Nat. Rev. Mol. Cell Biol.*, vol. 8, no. 7, pp. 562–573, 2007.

[37] R. Stephens, K. Lim, M. Portela, M. Kvansakul, P. O. Humbert, and H. E. Richardson, "The Scribble Cell Polarity Module in the Regulation of Cell Signaling in Tissue Development and Tumorigenesis," *J. Mol. Biol.*, pp. 1–28, 2018.

[38] M. Milgrom-Hoffman and P. O. Humbert, "Regulation of cellular and PCP signalling by the Scribble polarity module," *Semin. Cell Dev. Biol.*, 2017.

[39] S. Mili and I. G. Macara, "RNA localization and polarity: from A(PC) to Z(BP)," *Trends Cell Biol.*, vol. 19, no. 4, pp. 156–164, 2009.

[40] A. Román-Fernández and D. M. Bryant, "Complex Polarity: Building Multicellular Tissues Through Apical Membrane Traffic," *Traffic*, vol. 17, no. 12, pp. 1244–1261, 2016.

[41] T. Welz, J. Wellbourne-Wood, and E. Kerkhoff, "Orchestration of cell surface proteins by Rab11," *Trends Cell Biol.*, vol. 24, no. 7, pp. 407–414, 2014.

[42] K. VanderVorst, J. Hatakeyama, A. Berg, H. Lee, and K. L. Carraway, "Cellular and molecular mechanisms underlying planar cell polarity pathway contributions to cancer malignancy," *Semin. Cell Dev. Biol.*, pp. 1–10, 2017.

[43] M. P. Czech, "PIP2 and PIP3: complex roles at the cell surface.," *Cell*, vol. 100, no. 6, pp. 603–606, 2000.

[44] H. R. Pires and M. Boxem, "Mapping the Polarity Interactome," *J. Mol. Biol.*, 2018.

[45] S. Yogev and K. Shen, "Establishing Neuronal Polarity with Environmental and Intrinsic Mechanisms," *Neuron*, vol. 96, no. 3, pp. 638–650, 2017.

[46] A. P. Barnes and F. Polleux, "Establishment of Axon-Dendrite Polarity in Developing Neurons," *Annu. Rev. Neurosci.*, vol. 32, no. 1, pp. 347–381, 2009.

[47] P. Latzer, U. Schlegel, and C. Theiss, "Morphological Changes of Cortical and Hippocampal Neurons after Treatment with VEGF and Bevacizumab," *CNS Neurosci. Ther.*, vol. 22, no. 6, pp. 440–450, 2016.

[48] A. Sakakibara and Y. Hatanaka, "Neuronal polarization in the developing cerebral cortex," *Front. Neurosci.*, vol. 9, no. APR, pp. 1–10, 2015.

[49] M. P. Maya Shelly, Byung Kook Lim, Laura Cancedda, Sarah C. Heilshorn, Hongfeng Gao, “Local and Long-Range Reciprocal showed that localized cAMP and cGMP activ- ities were sufficient to induce preferential initia- tion of axons and dendrites, respectively (Fig. Regulation of cAMP and cGMP 1D). These effects are distinct from the known effe,” *Science (80-.).*, vol. 2, no. June, pp. 1–25, 2017.

[50] T. Takano *et al.*, “Discovery of long-range inhibitory signaling to ensure single axon formation,” *Nat. Commun.*, vol. 8, no. 1, pp. 1–17, 2017.

[51] Y. Mizoguchi and A. Monji, “Microglial Intracellular Ca2+ Signaling in Synaptic Development and its Alterations in Neurodevelopmental Disorders,” *Front. Cell. Neurosci.*, vol. 11, no. March, pp. 1–9, 2017.

[52] M. Schelski and F. Bradke, “Neuronal polarization: From spatiotemporal signaling to cytoskeletal dynamics,” *Mol. Cell. Neurosci.*, vol. 84, pp. 11–28, 2017.

[53] S. G. Tilson *et al.*, “ROCK inhibition facilitates in vitro expansion of glioblastoma stem-like cells,” *PLoS One*, vol. 10, no. 7, pp. 1–13, 2015.

[54] Y. Fujita and T. Yamashita, “Axon growth inhibition by RhoA/ROCK in the central nervous system,” *Front. Neurosci.*, vol. 8, no. OCT, pp. 1–12, 2014.

[55] E.-E. Govek *et al.*, “Cdc42 Regulates Neuronal Polarity during Cerebellar Axon Formation and Glial-Guided Migration,” *iScience*, pp. 35–48, 2018.

[56] E. K. Scott, J. E. Reuter, and L. Luo, “Small GTPase Cdc42 is required for multiple aspects of dendritic morphogenesis.,” *J. Neurosci.*, vol. 23, no. 8, pp. 3118–3123, 2003.

[57] G. Gallo, P. C. Letourneau, and C. A. M. Sam, “Axon guidance: GTPases help axons reach their targets.,” *Curr. Biol.*, vol. 8, no. 3, pp. R80-2, 1998.

[58] L. Dehmelt and S. Halpain, “Protein family review The MAP2 / Tau family of microtubule-associated proteins,” *Genome Biol.*, vol. 6, pp. 1–10, 2004.

[59] P. W. Baas, A. N. Rao, A. J. Matamoros, and L. Leo, “Stability properties of neuronal microtubules,” *Cytoskeleton*, vol. 73, no. 9, pp. 442–460, 2016.

[60] S. F. B. Van Beuningen *et al.*, “TRIM46 Controls Neuronal Polarity and Axon Specification by Driving the Formation of Parallel Microtubule Arrays,” *Neuron*, vol. 88, no. 6, pp. 1208–1226, 2015.

[61] A. N. Rao *et al.*, “Cytoplasmic Dynein Transports Axonal Microtubules in a Polarity-Sorting Manner,” *Cell Rep.*, vol. 19, no. 11, pp. 2210–2219, 2017.

[62] T. A. Maniar *et al.*, “UNC-33 (CRMP) and ankyrin organize microtubules and localize kinesin to polarize axon-dendrite sorting,” *Nat. Neurosci.*, vol. 15, no. 1, pp. 48–56, 2012.

[63] K. Xu, G. Zhong, and X. Zhuang, “Actin, Spectrin, and Associated Proteins Form a Periodic Cytoskeletal.pdf,” *Science*, vol. 339, no. January, pp. 452–6, 2013.

[64] S. L. Jones and T. M. Svitkina, “Axon Initial Segment Cytoskeleton: Architecture, Development, and Role in Neuron Polarity,” *Neural Plast.*, vol. 2016, 2016.

[65] P. D. Sarmiere and J. R. Bamburg, “Regulation of the Neuronal Actin Cytoskeleton by ADF/Cofilin,” *J. Neurobiol.*, vol. 58, no. 1, pp. 103–117, 2004.

[66] H. Kang and E. M. Schuman, “A Requirement for Local Protein Synthesis in Neurotrophin-Induced Hippocampal Synaptic Plasticity,” *Science (80-.).*, vol. 273, no. 5280, pp. 1402–1406, 1996.

[67] K. C. Martin *et al.*, “Synapse-specific, long-term facilitation of aplysia sensory to motor synapses: A function for local protein synthesis in memory storage,” *Cell*, vol. 91, no. 7, pp. 927–938, 1997.

[68] K. M. Huber, M. S. Kayser, and M. F. Bear, “Role for rapid dendritic protein synthesis in hippocampal mGluR- dependent long-term depression,” *Science (80-.).*, vol. 288, no. 5469, pp. 1254–1256, 2000.

[69] K. D. Bradshaw, N. J. Emptage, and T. V. P. Bliss, “A role for dendritic protein synthesis in hippocampal late LTP,” *Eur. J. Neurosci.*, vol. 18, no. 11, pp. 3150–3152, 2003.

[70] W. B. Smith, S. R. Starck, R. W. Roberts, and E. M. Schuman, “Dopaminergic stimulation of local protein synthesis enhances surface expression of GluR1 and synaptic transmission in hippocampal neurons,” *Neuron*, vol. 45, no. 5, pp. 765–779, 2005.

[71] A. G. M. Leenders and Z. H. Sheng, “Modulation of neurotransmitter release by the second messenger-activated protein kinases: Implications for presynaptic plasticity,” *Pharmacology and Therapeutics*, vol. 105, no. 1. pp. 69–84, 2005.

[72] M. A. Sutton and E. M. Schuman, “Dendritic Protein Synthesis, Synaptic Plasticity, and Memory,” *Cell*, vol. 127, no. 1. pp. 49–58, 2006.

[73] M. A. Sutton, “Intermediate-Term Memory for Site-Specific Sensitization in Aplysia Is Maintained by Persistent Activation of Protein Kinase C,” *J. Neurosci.*, vol. 24, no. 14, pp. 3600–3609, 2004.

[74] C. Rejon, M. Al-Masri, and L. McCaffrey, “Cell Polarity Proteins in Breast Cancer Progression,” *J. Cell. Biochem.*, vol. 2223, no. March, pp. 2215–2223, 2016.

[75] A. Gandalovičová, T. Vomastek, D. Rosel, and J. Brábek, “Cell polarity signaling in the plasticity of cancer cell invasiveness,” *Oncotarget*, vol. 5, no. 18, pp. 25022–49, 2016.

[76] F. Papagiannouli, “The internal structure of embryonic gonads and testis development in Drosophila melanogaster requires scrib, lgl and dlg activity in the soma,” *Int. J. Dev. Biol.*, vol. 57, no. 1, pp. 25–34, 2013.

[77] A. E. G. Idella F. Yamben, Rivka A. Rachel, Shalini Shetadal, Neal G. Copeland, Nancy A. Jenkins, Soren Warming, “Scrib is Required for Epithelial Cell Identity and Prevents Epithelial To Mesenchymal Transition in the Mouse,” *Dev. Biol.*, vol. 13, no. 2, pp. 83–96, 2013.

[78] T. H. Yair Adereth, Vincent Dammai, Nurgun Kose, Runzhao Li, “RNA-dependent integrin α3 protein localization regulated by the Muscleblind-like protein MLP1,” *Nat. Cell. Biol.*, vol. 7, no. 12, pp. 1240–1247, 2005.

[79] E. T. Wang *et al.*, “Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins,” *Cell*, vol. 150, no. 4, pp. 710–724, 2012.

[80] E. T. Wang *et al.*, “Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins,” *Genome Res.*, vol. 25, no. 6, pp. 858–871, 2015.

[81] J. M. Taliaferro *et al.*, “Distal Alternative Last Exons Localize mRNAs to Neural Projections,” *Mol. Cell*, vol. 61, no. 6, pp. 821–833, 2016.

[82] M. Warnstedt *et al.*, “The myotonia congenita mutation A331T confers a novel hyperpolarization-activated gate to the muscle chloride channel ClC-1.,” *J. Neurosci.*, vol. 22, no. 17, pp. 7462–7470, 2002.

[83] R. S. Savkur, A. V. Philips, and T. A. Cooper, “Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy,” *Nat. Genet.*, vol. 29, no. 1, pp. 40–47, 2001.

[84] C. Fugier *et al.*, “Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy,” *Nat. Med.*, vol. 17, no. 6, pp. 720–725, 2011.

[85] Z. Z. Tang *et al.*, “Muscle weakness in myotonic dystrophy associated with misregulated splicing and altered gating of Cav1.1 calcium channel,” *Hum. Mol. Genet.*, vol. 21, no. 6, pp. 1312–1324, 2012.

[86] C. A. Remme and C. R. Bezzina, “Sodium channel (Dys)function and cardiac arrhythmias,” *Cardiovasc. Ther.*, vol. 28, no. 5, pp. 287–294, 2010.

[87] M. M. Javadpour, J. C. Tardiff, I. Pinz, and J. S. Ingwall, “Decreased energetics in murine hearts bearing the R92Q mutation in cardiac troponin T,” *J. Clin. Invest.*, vol. 112, no. 5, pp. 768–775, 2003.

[88] S. Shukla and R. Parker, “Hypo- and Hyper-Assembly Diseases of RNA – Protein Complexes,” *Trends Mol. Med.*, vol. 22, no. 7, pp. 615–628, 2016.

[89] D. J. Battle *et al.*, “The SMN complex: An assembly machine for RNPs,” *Cold Spring Harb. Symp. Quant. Biol.*, vol. 71, pp. 313–320, 2006.

[90] C. Fallini, P. G. Donlin-Asp, J. P. Rouanet, G. J. Bassell, and W. Rossoll, “Deficiency of the Survival of Motor Neuron Protein Impairs mRNA Localization and Local Translation in the Growth Cone of Motor Neurons,” *J. Neurosci.*, vol. 36, no. 13, pp. 3811–3820, 2016.

[91] D. A. R. Zorio, C. M. Jackson, Y. Liu, E. W. Rubel, and Y. Wang, “Cellular distribution of the fragile X mental retardation protein in the mouse brain,” *J. Comp. Neurol.*, vol. 525, no. 4, pp. 818–849, 2017.

[92] J. A. Lee *et al.*, “Cytoplasmic Rbfox1 Regulates the Expression of Synaptic and Autism-Related Genes,” *Neuron*, vol. 89, no. 1, pp. 113–128, 2016.

[93] M. A. Kiebler and G. J. Bassell, “Neuronal RNA Granules: Movers and Makers,” *Neuron*, vol. 51, no. 6, pp. 685–690, 2006.

[94] M. Zeitelhofer, P. Macchi, and R. Dahm, “Perplexing bodies: The putative roles of P-bodies in neurons,” *RNA Biol.*, vol. 5, no. 4, pp. 244–248, 2008.

[95] A. M. Krichevsky and K. S. Kosik, “Neuronal RNA granules: A link between RNA localization and stimulation-dependent translation,” *Neuron*, vol. 32, no. 4, pp. 683–696, 2001.

[96] S. M. Fernandez-Moya, K. E. Bauer, and M. A. Kiebler, “Meet the players: local translation at the synapse,” *Front. Mol. Neurosci.*, vol. 7, no. November, pp. 1–6, 2014.

[97] C. J. Costa and D. E. Willis, “To the end of the line: Axonal mRNA transport and local translation in health and neurodegenerative disease,” *Developmental Neurobiology*, vol. 78, no. 3. pp. 209–220, 2018.

[98] E. T. Wang *et al.*, “Dysregulation of mRNA Localization and Translation in Genetic Disease,” *J. Neurosci.*, vol. 36, no. 45, pp. 11418–11426, 2016.

[99] T. Moss, “At the crossroads of growth control; making ribosomal RNA,” *Current Opinion in Genetics and Development*, vol. 14, no. 2. pp. 210–217, 2004.

[100] R. J. Sims, R. Belotserkovskaya, and D. Reinberg, “Elongation by RNA polymerase II : the short and long of it,” *Genes Dev.*, vol. 18, no. 20, pp. 2437–2468, 2004.

[101] G. Dieci, G. Fiorino, M. Castelnuovo, M. Teichmann, and A. Pagano, “The expanding RNA polymerase III transcriptome,” *Trends in Genetics*, vol. 23, no. 12. pp. 614–622, 2007.

[102] O. Kelemen *et al.*, “Function of alternative splicing,” *Gene*, vol. 514, no. 1. pp. 1–30, 2013.

[103] Z. Wang and C. B. Burge, “Splicing regulation: From a parts list of regulatory elements to an integrated splicing code,” *RNA*, vol. 14, no. 5, pp. 802–813, 2008.

[104] A. J. Matlin, F. Clark, and C. W. J. Smith, “Understanding alternative splicing: Towards a cellular code,” *Nature Reviews Molecular Cell Biology*, vol. 6, no. 5. pp. 386–398, 2005.

[105] S. M. Berget, C. Moore, and P. A. Sharp, “Spliced segments at the 5’ terminus of adenovirus 2 late mRNA,” *Proc. Natl. Acad. Sci.*, vol. 74, no. 8, pp. 3171–3175, 1977.

[106] L. T. Chow, R. E. Gelinas, T. R. Broker, and R. J. Roberts, “An amazing sequence arrangement at the 5’ ends of adenovirus 2 messenger RNA,” *Cell*, vol. 12, no. 1, pp. 1–8, 1977.

[107] R. F. Luco, M. Allo, I. E. Schor, A. R. Kornblihtt, and T. Misteli, “Epigenetics in alternative pre-mRNA splicing,” *Cell*, vol. 144, no. 1. pp. 16–26, 2011.

[108] G. Tushev, C. Glock, M. Heumüller, A. Biever, M. Jovanovic, and E. M. Schuman, “Alternative 3’ UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments,” *Neuron*, 2018.

[109] T. Shigeoka *et al.*, “Dynamic Axonal Translation in Developing and Mature Visual Circuits,” *Cell*, vol. 166, no. 1, pp. 181–192, 2016.

[110] J. M. Taliaferro *et al.*, “Distal Alternative Last Exons Localize mRNAs to Neural Projections,” *Mol. Cell*, 2016.

[111] Y. Katz, E. T. Wang, E. M. Airoldi, and C. B. Burge, “Analysis and design of RNA sequencing experiments for identifying isoform regulation,” *Nat. Methods*, vol. 7, no. 12, pp. 1009–1015, 2010.

[112] S. Lianoglou, V. Garg, J. L. Yang, C. S. Leslie, and C. Mayr, “Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression,” *Genes Dev.*, vol. 27, no. 21, pp. 2380–2396, 2013.

[113] M. Gu and C. D. Lima, “Processing the message: Structural insights into capping and decapping mRNA,” *Current Opinion in Structural Biology*, vol. 15, no. 1 SPEC. ISS. pp. 99–

106, 2005.

[114] J. Ross, “mRNA stability in mammalian cells.,” *Microbiol. Rev.*, vol. 59, no. 3, pp. 423–50, 1995.

[115] D. A. Mangus, M. C. Evans, and A. Jacobson, “Poly(A)-binding proteins: Multifunctional scaffolds for the post-transcriptional control of gene expression,” *Genome Biology*, vol. 4, no. 7. 2003.

[116] N. K. Gray, L. Hrabalkova, J. P. Scanlon, and R. W. P. Smith, “Poly(A)-binding proteins and mRNA localization: who rules the roost?,” *Biochem. Soc. Trans.*, vol. 43, no. 6, pp. 1277–1284, 2015.

[117] K. C. Martin and A. Ephrussi, “mRNA Localization: Gene Expression in the Spatial Dimension,” *Cell*, vol. 136, no. 4, pp. 719–730, 2009.

[118] C. Glock, M. Heumüller, and E. M. Schuman, “mRNA transport & local translation in neurons,” *Curr. Opin. Neurobiol.*, vol. 45, pp. 169–177, 2017.

[119] C. Eliscovich, A. R. Buxbaum, Z. B. Katz, and R. H. Singer, “mRNA on the move: The road to its biological destiny,” *J. Biol. Chem.*, vol. 288, no. 28, pp. 20361–20368, 2013.

[120] M. Feldbrügge, “mRNA transport meets membrane traffic,” *Trends Genet.*, vol. 30, no. 9, pp. 408–417, 2014.

[121] A. Chin and E. Lécuyer, “RNA localization: Making its way to the center stage,” *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1861, no. 11, pp. 2956–2970, 2017.

[122] M. W. Hentze, A. Castello, T. Schwarzl, and T. Preiss, “A brave new world of RNA-binding proteins,” *Nat. Rev. Mol. Cell Biol.*, 2018.

[123] J. G. Gall, “The centennial of the Cajal body,” *Nature Reviews Molecular Cell Biology*, vol. 4, no. 12. pp. 975–980, 2003.

[124] A. H. Fox *et al.*, “Paraspeckles: A novel nuclear domain,” *Curr. Biol.*, vol. 12, no. 1, pp. 13–25, 2002.

[125] M. Kulkarni, S. Ozgur, and G. Stoecklin, “On track with P-bodies,” *Biochem. Soc. Trans.*, vol. 38, no. 1, pp. 242–251, 2010.

[126] D. S. W. Protter and R. Parker, “Principles and Properties of Stress Granules,” *Trends in Cell Biology*, vol. 26, no. 9. pp. 668–679, 2016.

[127] N. Hirokawa, “mRNA Transport in Dendrites: RNA Granules, Motors, and Tracks,” *J. Neurosci.*, vol. 26, no. 27, pp. 7139–7142, 2006.

[128] M. Tolino, M. K??hrmann, and M. A. Kiebler, “RNA-binding proteins involved in RNA localization and their implications in neuronal diseases,” *Eur. J. Neurosci.*, vol. 35, no. 12, pp. 1818–1836, 2012.

[129] A. Castello, B. Fischer, M. W. Hentze, and T. Preiss, “RNA-binding proteins in Mendelian disease,” *Trends in Genetics*, vol. 29, no. 5. pp. 318–327, 2013.

[130] K. E. Lukong, K. Chang, E. W. Khandjian, and S. Richard, “RNA-binding proteins in human genetic disease.,” *Trends Genet.*, vol. 24, no. 8, pp. 416–425, 2008.

[131] P. G. Donlin-Asp *et al.*, “The Survival of Motor Neuron Protein Acts as a Molecular Chaperone for mRNP Assembly,” *Cell Rep.*, vol. 18, no. 7, pp. 1660–1673, 2017.

[132] R. H. Singer, “RNA zipcodes for cytoplasmic addresses,” *Curr. Biol.*, vol. 3, no. 10, pp. 719–721, 1993.

[133] E. H. Kislauskis, X. Zhu, and R. H. Singer, “Sequences responsible for intracellular localization of beta-actin messenger RNA also affect cell phenotype.,” *J. Cell Biol.*, vol. 127, no. 2, pp. 441–51, 1994.

[134] A. F. Ross, Y. Oleynikov, E. H. Kislauskis, K. L. Taneja, and R. H. Singer, “Characterization of a beta-actin mRNA zipcode-binding protein.,” *Mol. Cell. Biol.*, vol. 17, no. 4, pp. 2158–2165, 1997.

[135] Y. Shav-tal and R. H. Singer, “RNA localization,” *J Cell Sci.*, vol. 118, no. Pt 18, pp. 4077–4081, 2017.

[136] J.-M. Dahlene Fusco, Nathalie Accornero, Brigitte Lavoie, Shailesh M. Shenoy and and E. B. Blanchard, Robert H. Singer, “Single mRNA Molecules Demonstrate Probabilistic Movement in Living Mammalian Cells,” *Curr Biol.*, vol. 13, no. 2, pp. 161–167, 2003.

[137] E. Park and L. E. Maquat, “Staufen-mediated mRNA decay,” *Wiley Interdiscip Rev RNA*, vol. 4, no. 4, pp. 423–435, 2013.

[138] L. C. Kapitein and C. C. Hoogenraad, “Which way to go? Cytoskeletal organization and polarized transport in neurons,” *Mol. Cell. Neurosci.*, vol. 46, no. 1, pp. 9–20, 2011.

[139] M. A. Welte, “Bidirectional transport along microtubules,” *Curr. Biol.*, vol. 14, no. 13, pp. 525–537, 2004.

[140] K. M. F. and E. R. Gavis, “Live Imaging of Endogenous RNA Reveals a Diffusion and Entrapment Mechanism for nanos mRNA Localization in Drosophila,” *Curr. Biol.*, vol. 13, pp. 1159–1168, 2003.

[141] V. L. Zimyanin *et al.*, “In Vivo Imaging of oskar mRNA Transport Reveals the Mechanism of Posterior Localization,” *Cell*, vol. 134, no. 5, pp. 843–853, 2008.

[142] M. Kanke and P. M. Macdonald, “Translational activation of Oskar mRNA: Reevaluation of the role and importance of a 5′ regulatory element,” *PLoS One*, vol. 10, no. 5, pp. 1–16, 2015.

[143] V. Trovisco *et al.*, “bicoid mRNA localises to the Drosophila oocyte anterior by random Dynein-mediated transport and anchoring,” *Elife*, vol. 5, no. OCTOBER2016, pp. 1–34, 2016.

[144] and G. L. Lisa A. Mingle, Nataly N. Okuhama, Jian Shi1, Robert H. Singer, John Condeelis, “Localization of all seven messenger RNAs for the actin- polymerization nucleator Arp2/3 complex in the protrusions of fibroblasts,” *J Cell Sci.*, vol. 118, no. 11, pp. 2425–2433, 2005.

[145] R. Delanoue and I. Davis, “Dynein anchors its mRNA cargo after apical transport in the Drosophila blastoderm embryo,” *Cell*, vol. 122, no. 1, pp. 97–106, 2005.

[146] R. Delanoue, B. Herpers, J. Soetaert, I. Davis, and C. Rabouille, “Drosophila Squid/hnRNP Helps Dynein Switch from a gurken mRNA Transport Motor to an

Ultrastructural Static Anchor in Sponge Bodies," *Dev. Cell*, vol. 13, no. 4, pp. 523–538, 2007.

[147] D.-I. Kao, G. M. Aldridge, I. J. Weiler, and W. T. Greenough, "Altered mRNA transport, docking, and protein translation in neurons lacking fragile X mental retardation protein," *Proc. Natl. Acad. Sci.*, vol. 107, no. 35, pp. 15601–15606, 2010.

[148] J. B. Dictenberg, S. A. Swanger, L. N. Antar, R. H. Singer, and G. J. Bassell, "A Direct Role for FMRP in Activity-Dependent Dendritic mRNA Transport Links Filopodial-Spine Morphogenesis to Fragile X Syndrome," *Dev Cell.*, vol. 14, no. 6, pp. 926–939, 2008.

[149] A. R. Buxbaum, Y. J. Yoon, R. H. Singer, and H. Y. Park, "Single-molecule insights into mRNA dynamics in neurons," *Trends Cell Biol.*, vol. 25, no. 8, pp. 468–475, 2015.

[150] M. del R. Sánchez-Carbente and L. DesGroseillers, "Chapter 3 Understanding the importance of mRNA transport in memory," *Progress in Brain Research*, vol. 169. pp. 41–58, 2008.

[151] J. A. Ainsley, L. Drane, J. Jacobs, K. A. Kittelberger, and L. G. Reijmers, "Functionally diverse dendritic mRNAs rapidly associate with ribosomes following a novel experience," *Nat. Commun.*, vol. 5, 2014.

[152] E. Kim and H. Jung, "Local protein synthesis in neuronal axons: Why and how we study," *BMB Rep.*, vol. 48, no. 3, pp. 139–146, 2015.

[153] E. D. Pastuzyn *et al.*, "The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer," *Cell*, vol. 172, no. 1–2, p. 275–288.e18, 2018.

[154] M. Kato *et al.*, "Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels," *Cell*, vol. 149, no. 4, pp. 753–767, 2012.

[155] T. W. Han *et al.*, "Cell-free formation of RNA granules: Bound RNAs identify features and components of cellular assemblies," *Cell*, vol. 149, no. 4, pp. 768–779, 2012.

[156] C. T. Roberts and P. Kurre, "Vesicle trafficking and RNA transfer add complexity and connectivity to Cell-cell communication," *Cancer Res.*, vol. 73, no. 11, pp. 3200–3205, 2013.

[157] J. Mukherjee, B. Mahato, and S. Adhya, "Vesicular transport of a ribonucleoprotein to mitochondria," *Biol. Open*, vol. 3, no. 11, pp. 1083–1091, 2014.

[158] T. Janas, M. M. Janas, K. Sapoń, and T. Janas, "Mechanisms of RNA loading into exosomes," *FEBS Lett.*, vol. 589, no. 13, pp. 1391–1398, 2015.

[159] K. M. Kim, K. Abdelmohsen, M. Mustapic, D. Kapogiannis, and M. Gorospe, "RNA in extracellular vesicles," *Wiley Interdiscip. Rev. RNA*, vol. 8, no. 4, pp. 1–14, 2017.

[160] A. Alessia Deglincerti, Dilek Colak, Ulrich Hengst, Yaobin Liu, Guoqiang Xu and S. R. Jaffrey, "Coupled local translation and degradation regulate growth cone collapse," *Nat. Commun.*, vol. 6, pp. 69–81, 2015.

[161] K. Y. Wu *et al.*, "Local translation of RhoA regulates growth cone collapse," *Nature*, vol. 436, no. 7053, pp. 1020–1024, 2005.

[162] G. Gonsalvez and R. Long, "Spatial regulation of translation through RNA localization," *F1000 Biol. Rep.*, vol. 4, no. August, 2012.

[163] L. L. Chen and G. G. Carmichael, “Nuclear editing of mRNA 3b2-UTRs,” *Curr. Top. Microbiol. Immunol.*, vol. 353, no. 1, pp. 111–121, 2012.

[164] L. L. Chen, J. N. DeCerbo, and G. G. Carmichael, “Alu element-mediated gene silencing,” *EMBO J.*, vol. 27, no. 12, pp. 1694–1705, 2008.

[165] K. V. Prasanth *et al.*, “Regulating gene expression through RNA nuclear retention,” *Cell*, vol. 123, no. 2, pp. 249–263, 2005.

[166] C. R. Capshew, K. L. Dusenbury, and H. A. Hundley, “Inverted Alu dsRNA structures do not affect localization but can alter translation efficiency of human mRNAs independent of RNA editing,” *Nucleic Acids Res.*, vol. 40, no. 17, pp. 8637–8645, 2012.

[167] G. M. Borchert *et al.*, “Adenosine deamination in human transcripts generates novel microRNA binding sites,” *Hum. Mol. Genet.*, vol. 18, no. 24, pp. 4801–4807, 2009.

[168] H. F. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, and D. James, *Molecular Cell Biology*, vol. 5. 2008.

[169] N. Hug, D. Longman, and J. F. Cáceres, “Mechanism and regulation of the nonsense-mediated decay pathway,” *Nucleic Acids Res.*, vol. 44, no. 4, pp. 1483–1495, 2015.

[170] D. Colak, S. J. Ji, B. T. Porse, and S. R. Jaffrey, “Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay,” *Cell*, vol. 153, no. 6, 2013.

[171] B. J. Dickson and Y. Zou, “Navigating intermediate targets: the nervous system midline.,” *Cold Spring Harb. Perspect. Biol.*, vol. 2, no. 8, pp. 1–17, 2010.

[172] P. A. Brittis, Q. Lu, and J. G. Flanagan, “Axonal protein synthesis provides a mechanism for localized regulation at an intermediate target,” *Cell*, vol. 110, no. 2, pp. 223–235, 2002.

[173] Z. Chen, B. B. Gore, H. Long, L. Ma, and M. Tessier-Lavigne, “Alternative Splicing of the Robo3 Axon Guidance Receptor Governs the Midline Switch from Attraction to Repulsion,” *Neuron*, vol. 58, no. 3, pp. 325–332, 2008.

[174] D. L. Black and S. L. Zipursky, “To Cross or Not to Cross: Alternatively Spliced Forms of the Robo3 Receptor Regulate Discrete Steps in Axonal Midline Crossing,” *Neuron*, vol. 58, no. 3, pp. 297–298, 2008.

[175] S. Jonas and E. Izaurralde, “Towards a molecular understanding of microRNA-mediated gene silencing,” *Nature Reviews Genetics*. 2015.

[176] R. W. Carthew and E. J. Sontheimer, “Origins and Mechanisms of miRNAs and siRNAs,” *Cell*, vol. 136, no. 4, pp. 642–655, 2009.

[177] B. Czech and G. J. Hannon, “One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing,” *Trends Biochem. Sci.*, vol. 41, no. 4, pp. 324–337, 2016.

[178] S. Houwing *et al.*, “A Role for Piwi and piRNAs in Germ Cell Maintenance and Transposon Silencing in Zebrafish,” *Cell*, vol. 129, no. 1, pp. 69–82, 2007.

[179] V. V Vagin, A. Sigova, C. Li, V. Gvozdev, and P. D. Zamore, “A Distinct Small RNA Pathway Silences Selfish Genetic Elements in the Germline,” *Science (80-.).*, vol. 313, no. July, pp. 320–324, 2006.

[180] M. Lizio *et al.*, “Gateways to the FANTOM5 promoter level mammalian expression atlas,” *Genome Biol.*, vol. 16, no. 1, 2015.

[181] T. FANTOM Consortium and T. Riken Pmi, “A promoter-level mammalian expression atlas,” *Nature*, vol. 507, p. 462, 2014.

[182] Y. Fang and M. J. Fullwood, “Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer,” *Genomics, Proteomics Bioinforma.*, vol. 14, no. 1, pp. 42–54, 2016.

[183] F. Kopp and J. T. Mendell, “Functional Classification and Experimental Dissection of Long Noncoding RNAs,” *Cell*, vol. 172, no. 3, pp. 393–407, 2018.

[184] and M. G. Jesse M. Engreitz, Amy Pandya-Jones, Patrick McDonel, Alexander Shishkin, Klara Sirokman, Christine Surka, Sabah Kadri1, Jeffrey Xing, Alon Goren, Eric S. Lander, Kathrin Plath, “The Xist lncRNA exploits three-dimensional genome architecture to spread across the X-chromosome,” vol. 341, no. 6147, 2013.

[185] Arunoday Bhan and Subhrangsu S. Mandal, “LncRNA HOTAIR: a master regulator of chromatin dynamics and cancer,” vol. 1856, no. 1, pp. 151–164, 2015.

[186] Y. Lee, K. Jeon, J. T. Lee, S. Kim, and V. N. Kim, “MicroRNA maturation: Stepwise processing and subcellular localization,” *EMBO J.*, 2002.

[187] M. Chekulaeva and W. Filipowicz, “Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells,” *Current Opinion in Cell Biology*. 2009.

[188] W. Hu and J. Coller, “What comes first: Translational repression or mRNA degradation? the deepening mystery of microRNA function,” *Cell Res.*, vol. 22, no. 9, pp. 1322–1324, 2012.

[189] S. Sambandan *et al.*, “Activity-dependent spatially localized miRNA maturation in neuronal dendrites,” *Science (80-.).*, vol. 355, no. 6325, pp. 634–637, 2017.

[190] R. S. Muddashetty *et al.*, “Reversible Inhibition of PSD-95 mRNA Translation by miR-125a, FMRP Phosphorylation, and mGluR Signaling,” *Mol. Cell*, vol. 42, no. 5, pp. 673–688, 2011.

[191] F. Rehfeld *et al.*, “The RNA-binding protein ARPP21 controls dendritic branching by functionally opposing the miRNA it hosts,” *Nat. Commun.*, vol. 9, no. 1, 2018.

[192] S. Impey *et al.*, “An activity-induced microRNA controls dendritic spine formation by regulating Rac1-PAK signaling,” *Mol. Cell. Neurosci.*, vol. 43, no. 1, pp. 146–156, 2010.

[193] A. Bellon *et al.*, “miR-182 Regulates Slit2-Mediated Axon Guidance by Modulating the Local Translation of a Specific mRNA,” *Cell Rep.*, vol. 18, no. 5, pp. 1171–1186, 2017.

[194] T. L. Cuellar *et al.*, “Dicer loss in striatal neurons produces behavioral and neuroanatomical phenotypes in the absence of neurodegeneration,” *Proc. Natl. Acad. Sci.*, vol. 105, no. 14, pp. 5614–5619, 2008.

[195] G. M. Schratt *et al.*, “A brain-specific microRNA regulates dendritic spine development,” *Nature*, vol. 439, no. 7074, pp. 283–289, 2006.

[196] L. Han *et al.*, “Regulation of chemotropic guidance of nerve growth cones by microRNA,” *Mol. Brain*, vol. 4, no. 1, pp. 1–11, 2011.

[197] A. N. Iyer, A. Bellon, and M.-L. Baudet, “microRNAs in axon guidance,” *Front. Cell. Neurosci.*, vol. 8, no. March, pp. 1–14, 2014.

[198] U. Hengst, “Functional and Selective RNA Interference in Developing Axons and Growth Cones,” *J. Neurosci.*, vol. 26, no. 21, pp. 5727–5732, 2006.

[199] O. Natera-naranjo, A. Aschrafi, A. E. Gioio, and B. B. Kaplan, “Identification and quantitative analyses of microRNAs located in the distal axons of sympathetic neurons Identification and quantitative analyses of microRNAs located in the distal axons of sympathetic neurons,” *Spring*, pp. 1516–1529, 2010.

[200] and G. J. B. Yukio Sasaki, Christina Gross, Lei Xing, Yoshio Goshima, “Identification of Axon-Enriched MicroRNAs Localized to Growth Cones of Cortical Neurons,” *Dev Neurobiol.*, vol. 74, no. 3, pp. 397–406, 2014.

[201] M. C. Popis, S. Blanco, and M. Frye, “Post-transcriptional methylation of transfer and ribosomal RNA in stress response pathways , cell differentiation and cancer,” vol. 28, no. 1, pp. 65–71, 2016.

[202] G. Cao, H.-B. Li, Z. Yin, and R. A. Flavell, “Recent advances in dynamic m ^{6} A RNA modification,” *Open Biol.*, vol. 6, no. 4, p. 160003, 2016.

[203] J. Y. Roignant and M. Soller, “m6A in mRNA: An Ancient Mechanism for Fine-Tuning Gene Expression,” *Trends Genet.*, vol. 33, no. 6, pp. 380–390, 2017.

[204] K. D. Meyer *et al.*, “5’ UTR m6A Promotes Cap-Independent Translation,” *Cell*, vol. 163, no. 4, pp. 999–1010, 2015.

[205] D. Dominissini *et al.*, “Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq,” *Nature*, vol. 485, no. 7397, pp. 201–206, 2012.

[206] W. Xiao *et al.*, “Nuclear m 6 A Reader YTHDC1 Regulates mRNA Splicing,” *Mol. Cell*, vol. 61, no. 4, pp. 507–519, 2016.

[207] S. Ke *et al.*, “m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover,” *Genes Dev.*, vol. 31, no. 10, pp. 990–1006, 2017.

[208] J. Liu *et al.*, “VIRMA mediates preferential m6A mRNA methylation in 3’UTR and near stop codon and associates with alternative polyadenylation,” *Cell Discov.*, vol. 4, no. 1, 2018.

[209] H. Du *et al.*, “YTHDF2 destabilizes m 6 A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex,” *Nat. Commun.*, vol. 7, pp. 1–11, 2016.

[210] C. R. Alarcón, H. Lee, H. Goodarzi, N. Halberg, and S. F. Tavazoie, “N6 -methyl-adenosine (m6A) marks primary microRNAs for processing,” *Nature*, vol. 519, no. 7544, pp. 482–485, 2015.

[211] E. Martinez-Salas, R. Francisco-Velilla, J. Fernandez-Chamorro, and A. M. Embarek, “Insights into structural and mechanistic features of viral IRES elements,” *Front. Microbiol.*, vol. 8, no. JAN, pp. 1–15, 2018.

[212] H. Yamamoto, A. Unbehaun, and C. M. T. Spahn, “Ribosomal Chamber Music: Toward an Understanding of IRES Mechanisms,” *Trends Biochem. Sci.*, vol. 42, no. 8, pp. 655–668, 2017.

[213] M. A. Huynen, D. A. M. Konings, and P. Hogeweg, “Multiple coding and the evolutionary properties of RNA secondary structure,” *J. Theor. Biol.*, vol. 165, no. 2, pp. 251–267, 1993.

[214] J. G. Underwood *et al.*, “FragSeq: Transcriptome-wide RNA structure probing using high-throughput sequencing,” *Nat. Methods*, vol. 7, no. 12, pp. 995–1001, 2010.

[215] M. Kertesz *et al.*, “Genome-wide measurement of RNA secondary structure in yeast,” *Nature*, vol. 467, no. 7311, pp. 103–107, 2010.

[216] R. C. Spitale, P. Crisalli, R. A. Flynn, E. A. Torre, E. T. Kool, and H. Y. Chang, “RNA SHAPE analysis in living cells,” *Nat. Chem. Biol.*, vol. 9, no. 1, pp. 18–20, 2013.

[217] R. Soemedi *et al.*, “The effects of structure on pre-mRNA processing and stability,” *Methods*, vol. 125, pp. 36–44, 2017.

[218] N. N. Singh, R. N. Singh, and E. J. Androphy, “Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes,” *Nucleic Acids Res.*, vol. 35, no. 2, pp. 371–389, 2007.

[219] M. Kozak, “Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs.,” *Mol. Cell. Biol.*, vol. 9, no. 11, pp. 5134–5142, 1989.

[220] C. Pop *et al.*, “Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation,” *Mol. Syst. Biol.*, vol. 10, no. 12, pp. 770–770, 2014.

[221] J. R. Babendure, J. L. Babendure, J. Ding, and R. Y. Tsien, “Control of mammalian translation by mRNA structure near caps Control of mammalian translation by mRNA structure near caps,” *Rna*, vol. 12, no. Kozak 1994, pp. 851–861, 2006.

[222] B. D. Schneider and E. A. Leibold, “Effects of iron regulatory protein regulation on iron homeostasis during hypoxia,” *Regulation*, vol. 102, no. 9, pp. 3404–3411, 2003.

[223] A. V. Kochetov, A. Sarai, I. B. Rogozin, V. K. Shumny, and N. A. Kolchanov, “The role of alternative translation start sites in the generation of human protein diversity,” *Mol. Genet. Genomics*, 2005.

[224] D. R. Morris and A. P. Geballe, “Upstream Open Reading Frames as Regulators of mRNA Translation,” *Mol. Cell. Biol.*, 2000.

[225] G. L. Chew, A. Pauli, and A. F. Schier, “Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish,” *Nat. Commun.*, vol. 7, no. May, pp. 1–10, 2016.

[226] K. Wethmar, “The regulatory potential of upstream open reading frames in eukaryotic gene expression,” *Wiley Interdiscip. Rev. RNA*, vol. 5, no. 6, pp. 765–778, 2014.

[227] K. M. Vattem and R. C. Wek, “Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells,” *Proc. Natl. Acad. Sci.*, vol. 101, no. 31, pp. 11269–11274, 2004.

[228] J. Somers, T. Pöyry, and A. E. Willis, “A perspective on mammalian upstream open reading frame function,” *Int. J. Biochem. Cell Biol.*, vol. 45, no. 8, pp. 1690–1700, 2013.

[229] J. G. Gall and M. L. Pardue, “Formation and Detection of Rna-Dna Hybrid Molecules

in Cytological Preparations," *Proc. Natl. Acad. Sci.*, vol. 63, no. 2, pp. 378–383, 1969.

[230] A. M. Femino, F. S. Fay, K. Fogarty, and R. H. Singer, "Visualization of Single RNA Transcripts in Situ - Femino wt al. 1998," vol. 280, no. April, 1998.

[231] J. M. Halstead *et al.*, "An RNA biosensor for imaging the first round of translation from single cells to living animals," *Science (80-.).*, vol. 347, no. 6228, pp. 1367–1370, 2015.

[232] J. M. Levsky, "Fluorescence in situ hybridization: past, present and future," *J. Cell Sci.*, vol. 116, no. 14, pp. 2833–2838, 2003.

[233] F. Tubing, G. Vendra, M. Mikl, P. Macchi, S. Thomas, and M. A. Kiebler, "Dendritically Localized Transcripts Are Sorted into Distinct Ribonucleoprotein Particles That Display Fast Directional Motility along Dendrites of Hippocampal Neurons," *J. Neurosci.*, vol. 30, no. 11, pp. 4160–4170, 2010.

[234] V. Tatavarty *et al.*, "Single-molecule imaging of translational output from individual RNA granules in neurons," *Mol. Biol. Cell*, vol. 23, no. 5, pp. 918–929, 2012.

[235] E. Bertrand, P. Chartrand, M. Schaefer, S. M. Shenoy, R. H. Singer, and R. M. Long, "Localization of ASH1 mRNA particles in living yeast," *Mol. Cell*, vol. 2, no. 4, pp. 437–445, 1998.

[236] B. Wu, J. Chen, and R. H. Singer, "Background free imaging of single mRNAs in live cells using split fluorescent proteins," *Sci. Rep.*, vol. 4, pp. 11–13, 2014.

[237] M. O. Urbanek, P. Galka-Marciniak, M. Olejniczak, and W. J. Krzyzosiak, "RNA imaging in living cells – methods and applications," *RNA Biol.*, vol. 11, no. 8, pp. 1083–1095, 2014.

[238] B. C. Hebb and V. P. Whittaker, "INTRACELLULAR DISTRIBUTIONS OF ACETYLCHOLINE AND CHOLINE ACETYLASE," *J. Physiol.*, vol. 142, pp. 187–196, 1958.

[239] E. S. Lein *et al.*, "Genome-wide atlas of gene expression in the adult mouse brain," *Nature*, vol. 445, no. 7124, pp. 168–176, 2007.

[240] C. Trapnell, "Defining cell types and states with single-cell genomics," *Genome Research*. 2015.

[241] J. Zhong, T. Zhang, and L. M. Bloch, "Dendritic mRNAs encode diversified functionalities in hippocampal pyramidal neurons," *BMC Neurosci.*, vol. 7, pp. 1–12, 2006.

[242] M. M. Poon, S.-H. Choi, C. A. M. Jamieson, D. H. Geschwind, and K. C. Martin, "Identification of Process-Localized mRNAs from Cultured Rodent Hippocampal Neurons," *J. Neurosci.*, vol. 26, no. 51, pp. 13390–13399, 2006.

[243] R. F. Booth and J. B. Clark, "A rapid method for the preparation of relatively pure metabolically competent synaptosomes from rat brain.," *Biochem. J.*, vol. 176, no. 2, pp. 365–370, 1978.

[244] A. Nagy and A. V. Delgado-Escueta, "Rapid Preparation of Synaptosomes from Mammalian Brain Using Nontoxic Isoosmotic Gradient Material (Percoll)," *J. Neurochem.*, vol. 43, no. 4, pp. 1114–1123, 1984.

[245] G. J. O. Evans, "The synaptosome as a model system for studying synaptic

physiology," *Cold Spring Harb. Protoc.*, vol. 2015, no. 5, pp. 421–424, 2015.

[246] I. J. Cajigas, G. Tushev, T. J. Will, S. Tom Dieck, N. Fuerst, and E. M. Schuman, "The Local Transcriptome in the Synaptic Neuropil Revealed by Deep Sequencing and High-Resolution Imaging," *Neuron*, vol. 74, no. 3, pp. 453–466, 2012.

[247] S. V. Puthanveettil *et al.*, "A strategy to capture and characterize the synaptic transcriptome," *Proc. Natl. Acad. Sci.*, vol. 110, no. 18, pp. 7464–7469, 2013.

[248] B. J. Chen *et al.*, "RNA sequencing reveals pronounced changes in the noncoding transcriptome of aging synaptosomes," *Neurobiol. Aging*, vol. 56, pp. 67–77, 2017.

[249] T. J. Will *et al.*, "Deep sequencing and high-resolution imaging reveal compartment-specific localization of Bdnf mRNA in hippocampal neurons," *Sci. Signal.*, vol. 6, no. 306, 2013.

[250] D. Boyd-Kimball *et al.*, "Proteomic identification of proteins specifically oxidized by intracerebral injection of amyloid β-peptide (1-42) into rat brain: Implications for Alzheimer's disease," *Neuroscience*, vol. 132, no. 2, pp. 313–324, 2005.

[251] S. P. Schrimpf *et al.*, "Proteomic analysis of synaptosomes using isotope-coded affinity tags and mass spectrometry," *Proteomics*, vol. 5, no. 10, pp. 2531–2541, 2005.

[252] F. A. Witzmann *et al.*, "A proteomic survey of rat cerebral cortical synaptosomes," *Proteomics*, vol. 5, no. 8, pp. 2177–2201, 2006.

[253] J. D. Cahoy *et al.*, "A Transcriptome Database for Astrocytes, Neurons, and Oligodendrocytes: A New Resource for Understanding Brain Development and Function," *J. Neurosci.*, vol. 28, no. 1, pp. 264–278, 2008.

[254] B. W. Okaty, K. Sugino, and S. B. Nelson, "Cell Type-Specific Transcriptomics in the Brain," *J. Neurosci.*, vol. 31, no. 19, pp. 6939–6943, 2011.

[255] T. Klausberger and P. Somogyi, "Neuronal diversity and temporal dynamics: The unity of hippocampal circuit operations," *Science*, vol. 321, no. 5885. pp. 53–57, 2008.

[256] K. Sugino *et al.*, "Molecular taxonomy of major neuronal classes in the adult mouse forebrain," *Nat. Neurosci.*, vol. 9, no. 1, pp. 99–107, 2006.

[257] R. Daneman, L. Zhou, D. Agalliu, J. D. Cahoy, A. Kaushal, and B. A. Barres, "The mouse blood-brain barrier transcriptome: A new resource for understanding the development and function of brain endothelial cells," *PLoS One*, vol. 5, no. 10, 2010.

[258] S. Hanz *et al.*, "Axoplasmic importins enable retrograde injury signaling in lesioned nerve," *Neuron*, vol. 40, no. 6, pp. 1095–1104, 2003.

[259] L. J. Cox, U. Hengst, N. G. Gurskaya, K. A. Lukyanov, and S. R. Jaffrey, "Intra-axonal translation and retrograde trafficking of CREB promotes neuronal survival," *Nat. Cell Biol.*, vol. 10, no. 2, pp. 149–159, 2008.

[260] J. P. Junker *et al.*, "Genome-wide RNA Tomography in the Zebrafish Embryo," *Cell*, vol. 159, no. 3, pp. 662–675, 2014.

[261] D. O. Wang *et al.*, "Synapse- and stimulus-specific local translation during long-term neuronal plasticity," *Science (80-.).*, vol. 324, no. 5934, pp. 1536–1540, 2009.

[262] J. A. Troca-Marín, A. Alves-Sampaio, F. J. Tejedor, and M. L. Montesinos, “Local translation of dendritic RhoA revealed by an improved synaptoneurosome preparation,” *Mol. Cell. Neurosci.*, vol. 43, no. 3, pp. 308–314, 2010.

[263] T. Shigeoka, J. Jung, C. E. Holt, and H. Jung, “Axon-TRAP-RiboTag: Affinity purification of translated mrnas from neuronal axons in mouse in vivo,” in *Methods in Molecular Biology*, vol. 1649, 2018, pp. 85–94.

[264] M. Piper and C. Holt, “RNA TRANSLATION IN AXONS,” *Annu. Rev. Cell Dev. Biol.*, vol. 20, no. 1, pp. 505–523, 2004.

[265] E. Koenig, R. Martin, M. Titmus, and J. R. Sotelo-Silveira, “Cryptic peripheral ribosomal domains distributed intermittently along mammalian myelinated axons.,” *J. Neurosci.*, vol. 20, no. 22, pp. 390–400, 2000.

[266] A. Kun, L. Otero, J. R. Sotelo-Silveira, and J. R. Sotelo, “Ribosomal distributions in axons of mammalian myelinated fibers,” *J. Neurosci. Res.*, vol. 85, no. 10, pp. 2087–2098, 2007.

[267] B. A. Walker *et al.*, “Reprogramming axonal behavior by axon-specific viral transduction,” *Gene Ther.*, vol. 19, no. 9, pp. 947–955, 2012.

[268] M. H. G. Verheijen *et al.*, “Increased axonal ribosome numbers is an early event in the pathogenesis of amyotrophic lateral sclerosis,” *PLoS One*, vol. 9, no. 1, 2014.

[269] K. X. Zhang, L. Tan, M. Pellegrini, S. L. Zipursky, and J. M. McEwen, “Rapid Changes in the Translatome during the Conversion of Growth Cones to Synaptic Terminals,” *Cell Rep.*, vol. 14, no. 5, pp. 1258–1271, 2016.

[270] and M. L. Julie Lotharius, Jeppe Falsig, Johan van Beek, Sarah Payne, Ralf Dringen, Patrik Brundin, “Progressive Degeneration of Human Mesencephalic Neuron-Derived Cells Triggered by Dopamine-Dependent Oxidative Stress Is Dependent on the Mixed-Lineage Kinase Pathway,” *J. Neurosci.*, vol. 25, no. 27, pp. 6329–6342, 2005.

[271] D. Scholz *et al.*, “Rapid, complete and large-scale generation of post-mitotic neurons from the human LUHMES cell line,” *J. Neurochem.*, vol. 119, no. 5, pp. 957–971, 2011.

[272] P. Shastry, A. Basu, and M. S. Rajadhyaksha, “Neuroblastoma Cell Lines-A Versatile in Vztro Model in Neurobiology,” *Int. J. Neurosci.*, vol. 108, no. 42006, pp. 109–126, 2001.

[273] R. P. Castleberry, “Neuroblastoma,” vol. 33, no. 9, pp. 1430–1437, 1997.

[274] S. Påhlman *et al.*, “Differentiation and survival influences of growth factors in human neuroblastoma.,” *Eur. J. Cancer*, vol. 31A, no. 4, pp. 453–8, 1995.

[275] C. R. Nicholas and A. R. Kriegstein, “Cell reprogramming gets direct,” *Nature*, vol. 463, no. 7284, pp. 1031–1032, 2010.

[276] T. Vierbuchen, A. Ostermeier, Z. P. Pang, Y. Kokubu, T. C. Südhof, and M. Wernig, “Direct conversion of fibroblasts to functional neurons by defined factors,” *Nature*, vol. 463, no. 7284, pp. 1035–1041, 2010.

[277] M. L. Seibenhener and M. W. Wooten, “Isolation and Culture of Hippocampal Neurons from Prenatal Mice,” *J. Vis. Exp.*, no. 65, pp. 4–9, 2012.

[278] R. B. Campenot, “Local control of neurite development by nerve growth factor.,” *Proc. Natl. Acad. Sci. U. S. A.*, 1977.

[279] A. Aschrafi, A. N. Kar, O. Natera-Naranjo, M. A. MacGibeny, A. E. Gioio, and B. B. Kaplan, “MicroRNA-338 regulates the axonal expression of multiple nuclear-encoded mitochondrial mRNAs encoding subunits of the oxidative phosphorylation machinery,” *Cell. Mol. Life Sci.*, vol. 69, no. 23, pp. 4017–4027, 2012.

[280] O. E. Tasdemir-Yilmaz and R. A. Segal, “There and back again: Coordinated transcription, translation and transport in axonal survival and regeneration,” *Current Opinion in Neurobiology*. 2016.

[281] S. Boyden, “THE CHEMOTACTIC EFFECT OF MIXTURES OF ANTIBODY AND ANTIGEN ON POLYMORPHONUCLEAR LEUCOCYTES,” *J. Exp. Med.*, 1962.

[282] O. C. Pertz *et al.*, “Spatial mapping of the neurite and soma proteomes reveals a functional Cdc42/Rac regulatory network,” *Proc. Natl. Acad. Sci.*, vol. 105, no. 6, pp. 1931–1936, 2008.

[283] M. Iacovino, M. E. Roth, and M. Kyba, “Rapid genetic modification of mouse embryonic stem cells by inducible cassette exchange recombination,” *Methods Mol. Biol.*, vol. 1101, pp. 339–351, 2014.

[284] V. Yartseva, C. M. Takacs, C. E. Vejnar, M. T. Lee, and A. J. Giraldez, “RESA identifies mRNA-regulatory sequences at high resolution,” *Nat. Methods*, vol. 14, no. 2, pp. 201–207, 2017.

[285] J. E. Oh *et al.*, “Differentiation of neuroblastoma cell line N1E-115 involves several signaling cascades,” *Neurochem. Res.*, 2005.

[286] S. Clejan, R. S. Dotson, E. W. Wolf, M. P. Corb, and C. F. Ide, “Morphological differentiation of N1E-115 neuroblastoma cells by dimethyl sulfoxide activation of lipid second messengers,” *Exp. Cell Res.*, 1996.

[287] Y. Yakubchyk *et al.*, “Regulation of Neurite Outgrowth in N1E-115 Cells through PDZ-Mediated Recruitment of Diacylglycerol Kinase ,” *Mol. Cell. Biol.*, 2005.

[288] J. M. Rodrigues *et al.*, “Intracellular Ca 2+ concentration in the N1E-115 neuronal cell line and its use for peripheric nerve regeneration,” *Acta Med. Port.*, 2005.

[289] J. E. Oh, K. Karlmark Raja, J. H. Shin, A. Pollak, M. Hengstschläger, and G. Lubec, “Cytoskeleton changes following differentiation of N1E-115 neuroblastoma cell line,” in *Amino Acids*, 2006.

[290] H. Mi, A. Muruganujan, J. T. Casagrande, and P. D. Thomas, “Large-scale gene function analysis with the panther classification system,” *Nat. Protoc.*, vol. 8, no. 8, pp. 1551–1566, 2013.

[291] S. Chanda *et al.*, “Generation of induced neuronal cells by the single reprogramming factor ASCL1,” *Stem Cell Reports*, vol. 3, no. 2, pp. 282–296, 2014.

[292] B. Treutlein *et al.*, “Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq,” *Nature*, vol. 534, no. 7607, pp. 391–395, 2016.

[293] S.-E. Ong *et al.*, “Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a

Simple and Accurate Approach to Expression Proteomics," *Mol. Cell. Proteomics*, vol. 1, no. 5, pp. 376–386, 2002.

[294] S. Y. Cho and R. L. Klemke, "Purification of pseudopodia from polarized cells reveals redistribution and activation of Rac through assembly of a CAS/Crk scaffold," *J. Cell Biol.*, vol. 156, no. 4, pp. 725–736, 2002.

[295] F. K. Mardakheh *et al.*, "Global Analysis of mRNA, Translation, and Protein Localization: Local Translation Is a Key Regulator of Cell Protrusions," *Dev. Cell*, vol. 35, no. 3, pp. 344–357, 2015.

[296] M. Jarocki, O. Sallouh, R. Weberskirch, and A. Faissner, "The Tenascin-C-Derived Peptide VSWRAPTA Promotes Neuronal Branching Via Transcellular Activation of the Focal Adhesion Kinase (FAK) and the ERK1/2 Signaling Pathway In Vitro," *Molecular Neurobiology*, 2018.

[297] K. C. Hart, S. C. Robertson, M. Y. Kanemitsu, A. N. Meyer, J. A. Tynan, and D. J. Donoghue, "Transformation and Stat activation by derivatives of FGFR1, FGFR3, and FGFR4," *Oncogene*, 2000.

[298] H. Yang, H. Wang, Y. Shu, and X. Li, "miR-103 Promotes Neurite Outgrowth and Suppresses Cells Apoptosis by Targeting Prostaglandin-Endoperoxide Synthase 2 in Cellular Models of Alzheimer's Disease.," *Front. Cell. Neurosci.*, 2018.

[299] L. F. Lareau, R. E. Green, R. S. Bhatnagar, and S. E. Brenner, "The evolving roles of alternative splicing," *Current Opinion in Structural Biology*, vol. 14, no. 3. pp. 273–282, 2004.

[300] H. Gruner, M. Cortï¿½s-Lï¿½pez, D. A. Cooper, M. Bauer, and P. Miura, "CircRNA accumulation in the aging mouse brain," *Sci. Rep.*, vol. 6, 2016.

[301] M. Cortés-López and P. Miura, "Emerging functions of circular RNAs," *Yale Journal of Biology and Medicine*, vol. 89, no. 4. pp. 527–537, 2016.

[302] N. R. Pamudurti *et al.*, "Translation of CircRNAs," *Mol. Cell*, vol. 66, no. 1, p. 9–21.e7, 2017.

[303] T. B. Hansen *et al.*, "Natural RNA circles function as efficient microRNA sponges," *Nature*, vol. 495, no. 7441, pp. 384–388, 2013.

[304] S. Memczak *et al.*, "Circular RNAs are a large class of animal RNAs with regulatory potency," *Nature*, vol. 495, no. 7441, pp. 333–338, 2013.

[305] K. S. Cramer, D. P. Cerretti, and S. A. Siddiqui, "EphB2 regulates axonal growth at the midline in the developing auditory brainstem," *Dev. Biol.*, vol. 295, no. 1, pp. 76–89, 2006.

[306] Z. Quan, D. Zheng, and H. Qing, "Regulatory Roles of Long Non-Coding RNAs in the Central Nervous System and Associated Neurodegenerative Diseases," *Front. Cell. Neurosci.*, vol. 11, 2017.

[307] A. Zappulo *et al.*, "RNA localization is a key determinant of neurite-enriched proteome," *Nat. Commun.*, vol. 8, no. 1, 2017.

[308] R. B. Weinberg, E. J. Mufson, and S. E. Counts, "Evidence for a neuroprotective microRNA pathway in amnestic mild cognitive impairment," *Front. Neurosci.*, 2015.

[309] D. Most, C. Leiter, Y. A. Blednov, R. A. Harris, and R. D. Mayfield, “Synaptic micrornas coordinately regulate synaptic mRnas: Perturbation by chronic alcohol consumption,” *Neuropsychopharmacology*, 2016.

[310] P. Shu *et al.*, “MicroRNA-214 modulates neural progenitor cell differentiation by targeting Quaking during cerebral cortex development,” *Sci. Rep.*, 2017.

[311] F. Cimadamore, A. Amador-Arjona, C. Chen, C.-T. Huang, and A. V. Terskikh, “SOX2-LIN28/let-7 pathway regulates proliferation and neurogenesis in neural precursors,” *Proc. Natl. Acad. Sci.*, 2013.

[312] A. Selvamani, P. Sathyan, R. C. Miranda, and F. Sohrabji, “An antagomir to microRNA let7f promotes neuroprotection in an ischemic stroke model,” *PLoS One*, 2012.

[313] C. H. Chou *et al.*, “MiRTarBase update 2018: A resource for experimentally validated microRNA-target interactions,” *Nucleic Acids Res.*, vol. 46, no. D1, pp. D296–D302, 2018.

[314] S. Gerstberger, M. Hafner, and T. Tuschl, “A census of human RNA-binding proteins,” *Nat. Rev. Genet.*, vol. 15, no. 12, pp. 829–845, 2014.

[315] T. L. Bailey *et al.*, “MEME Suite: Tools for motif discovery and searching,” *Nucleic Acids Res.*, vol. 37, no. SUPPL. 2, 2009.

[316] T. L. Bailey and M. Gribskov, “Combining evidence using p-values: Application to sequence homology searches,” *Bioinformatics*, vol. 14, no. 1, pp. 48–54, 1998.

[317] D. Dominissini *et al.*, “The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA,” *Nature*, vol. 530, no. 7591, pp. 441–446, 2016.

[318] K. D. Meyer, Y. Saletore, P. Zumbo, O. Elemento, C. E. Mason, and S. R. Jaffrey, “Comprehensive analysis of mRNA methylation reveals enrichment in 3’ UTRs and near stop codons,” *Cell*, vol. 149, no. 7, pp. 1635–1646, 2012.

[319] Y. Wang, J. Wang, L. Gao, S. Stamm, and A. Andreadis, “An SRp75/hnRNPG complex interacting with hnRNPE2 regulates the 5’ splice site of tau exon 10, whose misregulation causes frontotemporal dementia,” *Gene*, 2011.

[320] and H. Y. C. John L. Rinn, Michael Kertesz, Jordon K. Wang, Sharon L. Squazzo, Xiao Xu, Samantha A. Brugmann, Henry Goodnough, Jill A. Helms, Peggy J. Farnham, Eran Segal, “Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Non-Coding RNAs,” *Cell.*, vol. 129, no. 7, pp. 1311–1323, 2007.

[321] K. B. Jensen and R. B. Darnell, “CLIP: Crosslinking and ImmunoPrecipitation of In Vivo RNA Targets of RNA-Binding Proteins,” vol. 488, pp. 85–98, 2008.

[322] T. Koressaar and M. Remm, “Enhancements and modifications of primer design program Primer3,” *Bioinformatics*, vol. 23, no. 10, pp. 1289–1291, 2007.

[323] A. Untergasser *et al.*, “Primer3-new capabilities and interfaces,” *Nucleic Acids Res.*, vol. 40, no. 15, 2012.

[324] Y. Zhang *et al.*, “Model-based analysis of ChIP-Seq (MACS),” *Genome Biol.*, vol. 9, no. 9, 2008.

[325] M. I. Love, S. Anders, and W. Huber, *Differential analysis of count data - the DESeq2*

package, vol. 15, no. 12. 2014.

[326] O. L. Wapinski *et al.*, "XHierarchical mechanisms for direct reprogramming of fibroblasts to neurons," *Cell*, 2013.

[327] M. Hausser, "The Beat Goes On: Spontaneous Firing in Mammalian Neuronal Microcircuits," *J. Neurosci.*, 2004.

[328] B. E. Pfeiffer and K. M. Huber, "Current Advances in Local Protein Synthesis and Synaptic Plasticity," *J. Neurosci.*, 2006.

[329] E. L. Spaulding and R. W. Burgess, "Accumulating evidence for axonal translation in neuronal homeostasis," *Frontiers in Neuroscience*. 2017.

[330] J. Sebeo *et al.*, "Requirement for Protein Synthesis at Developing Synapses," *J. Neurosci.*, 2009.

[331] C. Medioni, K. Mowry, and F. Besse, "Principles and roles of mRNA localization in animal development," *Development*, 2012.

[332] G. Liao, L. Mingle, L. Van De Water, and G. Liu, "Control of cell migration through mRNA localization and local translation," *Wiley Interdisciplinary Reviews: RNA*. 2015.

[333] I. Kronja and T. L. Orr-Weaver, "Translational regulation of the cell cycle: when, where, how and why?," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 2011.

[334] J. Barr, K. V. Yakovlev, Y. Shidlovskii, and P. Schedl, "Establishing and maintaining cell polarity with mRNA localization in Drosophila," *BioEssays*. 2016.

[335] D. S. Campbell and C. E. Holt, "Chemotropic responses of retinal growth cones mediated by rapid local protein synthesis and degradation," *Neuron*, vol. 32, no. 6, pp. 1013–1026, 2001.

[336] M. Spillane, A. Ketschek, T. T. Merianda, J. L. Twiss, and G. Gallo, "Mitochondria Coordinate Sites of Axon Branching through Localized Intra-axonal Protein Synthesis," *Cell Rep.*, vol. 5, no. 6, pp. 1564–1575, 2013.

[337] C. Bagni and W. T. Greenough, "From mRNP trafficking to spine dysmorphogenesis: The roots of fragile X syndrome," *Nature Reviews Neuroscience*, vol. 6, no. 5. pp. 376–387, 2005.

[338] J. C. Darnell and E. Klann, "The translation of translational control by FMRP: Therapeutic targets for FXS," *Nature Neuroscience*, vol. 16, no. 11. pp. 1530–1536, 2013.

[339] J. D. Richter, E. Klann, J. D. Richter, and E. Klann, "Making synaptic plasticity and memory last: mechanisms of translational regulation.," *Genes Dev.*, vol. 23, no. 1, pp. 1–11, 2009.

[340] Y.-S. Huang and J. D. Richter, "Analysis of mRNA translation in cultured hippocampal neurons.," *Methods Enzymol.*, vol. 431, no. 07, pp. 143–162, 2007.

[341] S. J. Ji and S. R. Jaffrey, "Intra-axonal Translation of SMAD1/5/8 Mediates Retrograde Regulation of Trigeminal Ganglia Subtype Specification," *Neuron*, vol. 74, no. 1, pp. 95–107, 2012.

[342] K. Ben-Yaakov *et al.*, "Axonal transcription factors signal retrogradely in lesioned

peripheral nerve," *EMBO J.*, vol. 31, no. 6, pp. 1350–1363, 2012.

[343] C. Andreassi *et al.*, "An NGF-responsive element targets myo-inositol monophosphatase-1 mRNA to sympathetic neuron axons," *Nat. Neurosci.*, vol. 13, no. 3, pp. 291–301, 2010.

[344] A. C. Lin and C. E. Holt, "Function and regulation of local axonal translation," *Current Opinion in Neurobiology*, vol. 18, no. 1. pp. 60–68, 2008.

[345] R. Narayanan and D. Johnston, "The h Channel Mediates Location Dependence and Plasticity of Intrinsic Phase Response in Rat Hippocampal Neurons," *J. Neurosci.*, vol. 28, no. 22, pp. 5846–5860, 2008.

[346] J. K. Makara, A. Losonczy, Q. Wen, and J. C. Magee, "Experience-dependent compartmentalized dendritic plasticity in rat hippocampal CA1 pyramidal neurons," *Nat. Neurosci.*, vol. 12, no. 12, pp. 1485–1487, 2009.

[347] K. F. Raab-graham, P. C. G. Haddick, Y. N. Jan, and L. Y. Jan, "Activity- and mTOR-Dependent Suppression of Kv1.1 Channel mRNA Translation in Dendrites," *Science (80-.).*, no. October, pp. 144–148, 2006.

[348] a Gardiol, C. Racca, and a Triller, "Dendritic and postsynaptic protein synthetic machinery.," *J. Neurosci.*, vol. 19, no. 1, pp. 168–179, 1999.

[349] A. C. Horton, B. Rácz, E. E. Monson, A. L. Lin, R. J. Weinberg, and M. D. Ehlers, "Polarized secretory trafficking directs cargo for asymmetric dendrite growth and morphogenesis," *Neuron*, vol. 48, no. 5, pp. 757–771, 2005.

[350] A. C. Horton and M. D. Ehlers, "Dual modes of endoplasmic reticulum-to-Golgi transport in dendrites revealed by live-cell imaging.," *J. Neurosci.*, vol. 23, no. 15, pp. 6188–6199, 2003.

[351] E. R. Torre and O. Steward, "Protein Synthesis within Dendrites: Glycosylation of Newly Synthesized Proteins in Dendrites of Hippocampal Neurons in Culture," *J. Neurosci.*, vol. 16, no. 19, pp. 5967–5978, 1996.

[352] H. Eng, K. Lund, and R. B. Campenot, "Synthesis of beta-tubulin, actin, and other proteins in axons of sympathetic neurons in compartmented cultures.," *J. Neurosci.*, vol. 19, no. 1, pp. 1–9, 1999.

[353] K. M. Leung, F. P. G. Van Horck, A. C. Lin, R. Allison, N. Standart, and C. E. Holt, "Asymmetrical β-actin mRNA translation in growth cones mediates attractive turning to netrin-1," *Nat. Neurosci.*, vol. 9, no. 10, pp. 1247–1256, 2006.

[354] J. Yao, Y. Sasaki, Z. Wen, G. J. Bassell, and J. Q. Zheng, "An essential role for beta-actin mRNA localization and translation in Ca2+-dependent growth cone guidance.," *Nat. Neurosci.*, vol. 9, no. 10, pp. 1265–1273, 2006.

[355] U. Hengst, A. Deglincerti, H. J. Kim, N. L. Jeon, and S. R. Jaffrey, "Axonal elongation triggered by stimulus-induced local translation of a polarity complex protein," *Nat. Cell Biol.*, vol. 11, no. 8, pp. 1024–1030, 2009.

[356] M. Piper *et al.*, "Signaling mechanisms underlying Slit2-induced collapse of Xenopus retinal growth cones," *Neuron*, vol. 49, no. 2, pp. 215–228, 2006.

[357] A. F. R. Batista and U. Hengst, “Intra-axonal protein synthesis in development and beyond,” *International Journal of Developmental Neuroscience*, vol. 55. pp. 140–149, 2016.

[358] K. Kalil and E. W. Dent, “Branch management: Mechanisms of axon branching in the developing vertebrate CNS,” *Nature Reviews Neuroscience*, vol. 15, no. 1. pp. 7–18, 2014.

[359] M. Spillane, A. Ketschek, C. J. Donnelly, A. Pacheco, J. L. Twiss, and G. Gallo, “Nerve Growth Factor-Induced Formation of Axonal Filopodia and Collateral Branches Involves the Intra-Axonal Synthesis of Regulators of the Actin-Nucleating Arp2/3 Complex,” *J. Neurosci.*, vol. 32, no. 49, pp. 17671–17689, 2012.

[360] C. J. Donnelly *et al.*, “Axonally Synthesized -Actin and GAP-43 Proteins Support Distinct Modes of Axonal Growth,” *J. Neurosci.*, vol. 33, no. 8, pp. 3311–3322, 2013.

[361] H. H. W. Wong *et al.*, “RNA Docking and Local Translation Regulate Site-Specific Axon Remodeling In Vivo,” *Neuron*, vol. 95, no. 4, p. 852–868.e8, 2017.

[362] A. F. R. Batista, J. C. Martínez, and U. Hengst, “Intra-axonal Synthesis of SNAP25 Is Required for the Formation of Presynaptic Terminals,” *Cell Rep.*, vol. 20, no. 13, pp. 3085–3098, 2017.

[363] A. M. Taylor, J. Wu, H. C. Tai, and E. M. Schuman, “Axonal translation of beta-catenin regulates synaptic vesicle dynamics,” *J. Neurosci.*, vol. 33, no. 13, pp. 5584–5589, 2013.

[364] K. Hsiao, O. Bozdagi, and D. L. Benson, “Axonal cap-dependent translation regulates presynaptic p35,” *Dev. Neurobiol.*, vol. 74, no. 3, pp. 351–364, 2014.

[365] H. Hörnberg and C. Holt, “RNA-binding proteins and translational regulation in axons and growth cones,” *Front. Neurosci.*, vol. 7, no. 7 MAY, pp. 1–9, 2013.

[366] H. Jung, B. C. Yoon, and C. E. Holt, “Axonal mRNA localization and local protein synthesis in nervous system assembly, maintenance and repair,” *Nat. Rev. Neurosci.*, vol. 13, no. 5, pp. 308–324, 2012.

[367] H. Jung, C. G. Gkogkas, N. Sonenberg, and C. E. Holt, “Remote control of gene function by local translation,” *Cell*, vol. 157, no. 1, pp. 26–40, 2014.

[368] N. Sonenberg and A. G. Hinnebusch, “Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets,” *Cell*, vol. 136, no. 4, pp. 731–745, 2009.

[369] P. K. Arthur, M. Claussen, S. Koch, K. Tarbashevich, O. Jahn, and T. Pieler, “Participation of Xenopus Elr-type proteins in vegetal mRNA localization during oogenesis,” *J. Biol. Chem.*, 2009.

[370] S. Berto, N. Usui, G. Konopka, and B. L. Fogel, “ELAVL2-regulated transcriptional and splicing networks in human neurons link neurodevelopment and autism,” *Hum. Mol. Genet.*, 2016.

[371] J. Katahira *et al.*, “Nuclear RNA export factor 7 is localized in processing bodies and neuronal RNA granules through interactions with shuttling hnRNPs,” *Nucleic Acids Res.*, 2008.

[372] A. Lux *et al.*, “Human retroviral gag- and gag-pol-like proteins interact with the transforming growth factor-β receptor activin receptor-like kinase 1,” *J. Biol. Chem.*, 2005.

[373] S. Xue and M. Barna, “Specialized ribosomes: A new frontier in gene regulation and organismal biology,” *Nature Reviews Molecular Cell Biology*. 2012.

[374] J. Pohlenz *et al.*, “Partial deficiency of thyroid transcription factor 1 produces predominantly neurological defects in humans and mice,” *J. Clin. Invest.*, 2002.

[375] C. Yuan, P. Li, S. Guo, B. Zhang, T. Sun, and J. Cui, “The Role of Purα in Neuronal Development, the Progress in the Current Researches,” *J. Neurol. Neurosci.*, vol. 07, no. 05, pp. 1–6, 2016.

[376] B. Wang and L. Bao, “Axonal microRNAs: Localization, function and regulatory mechanism during axon development,” *J. Mol. Cell Biol.*, vol. 9, no. 2, pp. 82–90, 2017.

[377] G. M. Davis, M. A. Haas, and R. Pocock, “MicroRNAs: Not ‘Fine-Tuners’ but key regulators of neuronal development and function,” *Front. Neurol.*, vol. 6, no. NOV, pp. 1–12, 2015.

[378] J. Yu *et al.*, “Dynamic m6A modification regulates local translation of mRNA in axons,” *Nucleic Acids Res.*, vol. 46, no. 3, pp. 1412–1423, 2018.

[379] J. Tcherkezian, P. A. Brittis, F. Thomas, P. P. Roux, and J. G. Flanagan, “Transmembrane Receptor DCC Associates with Protein Synthesis Machinery and Regulates Translation,” *Cell*, vol. 141, no. 4, pp. 632–644, 2010.

[380] S. Meng *et al.*, “CircRNA: Functions and properties of a novel potential biomarker for cancer,” *Molecular Cancer*. 2017.

[381] A. Rybak-Wolf *et al.*, “Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed,” *Mol. Cell*, 2014.

[382] C. Michaloglou *et al.*, “The Tyrosine Phosphatase PTPN14 Is a Negative Regulator of YAP Activity,” *PLoS One*, 2013.

[383] A. Hergovich, “MOB control: Reviewing a conserved family of kinase regulators,” *Cellular Signalling*. 2011.

[384] C. D. Bryan, C. Bin Chien, and K. M. Kwan, “Loss of laminin alpha 1 results in multiple structural defects and divergent effects on adhesion during vertebrate optic cup morphogenesis,” *Dev. Biol.*, 2016.

[385] K. L. Obholz, A. Akopyan, K. G. Waymire, and G. R. MacGregor, “FNDC3A is required for adhesion between spermatids and Sertoli cells,” *Dev. Biol.*, 2006.

[386] R. Gassmann *et al.*, “A new mechanism controlling kinetochore-microtubule interactions revealed by comparison of two dynein-targeting components: SPDL-1 and the Rod/Zwilch/Zw10 complex,” *Genes Dev.*, 2008.

[387] F. F. J. Kersten *et al.*, “The mitotic spindle protein SPAG5/Astrin connects to the Usher protein network postmitotically,” *Cilia*, 2012.